

## **Title of paper:**

Comprehensive genome sequencing analysis as a promising option in the prenatal diagnosis of fetal structural anomalies: a prospective study.

## **Authors' full name and academic degrees:**

Jia Zhou, MD<sup>1</sup>, Ziying Yang, MS<sup>2,3</sup>, Jun Sun, PHD<sup>2,3</sup>, Lipei Liu, PHD<sup>2,3</sup>, Xinyao Zhou, PHD<sup>1</sup>, Fengxia Liu, MS<sup>2,3</sup>, Ya Xing, MD<sup>1</sup>, Shuge Cui, MS<sup>2,3</sup>, Shiyi Xiong, MD, PHD<sup>1</sup>, Xiaoyu Liu, MS<sup>2,3</sup>, Yingjun Yang, MD<sup>1</sup>, Xiuxiu Wei, MS<sup>2,3</sup>, Gang Zou, MD<sup>1</sup>, Zhonghua Wang, PHD<sup>2,3</sup>, Xing Wei, MD<sup>1</sup>, Yaoshen Wang, BS<sup>2,3</sup>, Yun Zhang, MD<sup>1</sup>, Saiying Yan, MS<sup>2,3</sup>, Fengyu Wu, MD<sup>1</sup>, Fanwei Zeng, MS<sup>2,5</sup>, Tao Duan, MD<sup>1</sup>, Jian Wang, PHD<sup>4</sup>, Yaping Yang, PHD<sup>6</sup>, Zhiyu Peng, PHD<sup>2</sup>, Luming Sun, MD, PHD<sup>1</sup>

## **Authors' primary affiliations:**

<sup>1</sup> Shanghai First Maternity and Infant Hospital, Tongji University School of Medicine, Shanghai, China;

<sup>2</sup> BGI Genomics, BGI-Shenzhen, Shenzhen, China;

<sup>3</sup> Tianjin Medical Laboratory, BGI-Tianjin, BGI-Shenzhen, Tianjin, China;

<sup>4</sup> Department of Medical Genetics and Molecular Diagnostic Laboratory, Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine.

<sup>5</sup> Department of Biology, Faculty of Science, University of Copenhagen, Copenhagen, DK-2200, Denmark;

<sup>6</sup> AiLife Diagnostics, Pearland, TX 77584, USA.

**Correspondence:** Zhiyu Peng ([pengzhiyu@bgi.com](mailto:pengzhiyu@bgi.com)) and Luming Sun ([luming\\_sun@163.com](mailto:luming_sun@163.com))

These authors contributed equally: Jia Zhou, Ziying Yang and Jun Sun

## ABSTRACT

**Purpose:** Genome sequencing (GS) is a powerful tool for postnatal genetic diagnosis, but relevant clinical studies in the field of prenatal diagnosis are few. We aimed to evaluate the feasibility of GS as a first-line approach in prenatal diagnosis and compare its clinical value with the chromosomal microarray analysis (CMA) plus exome sequencing (ES) sequential testing.

**Methods:** We applied trio GS (~40-fold) in parallel with CMA plus ES to investigate the genetic basis for structural or growth anomalies in 111 fetuses and compared their results.

**Results:** GS covered all genetic variants in 22 diagnosed cases detected by CMA plus ES, yielding a diagnostic rate of 19.8% (22/110). Moreover, GS provided more comprehensive and precise genetic information than CMA plus ES, revealing twin fetuses with an imbalanced translocation arising from a balanced paternal translocation and one fetus with an extra pathogenic variant in the *GJA8* gene, and incidentally identified intrauterine CMV infection in a growth-restricted fetus.

**Conclusion:** Compared with CMA plus ES, GS offers a more comprehensive view of the genetic etiology of fetal anomalies and provides clues for nongenetic factors such as intrauterine infection. Our study demonstrates the feasibility of GS as a promising first-line test in prenatal diagnosis.

**Keywords:** Genome sequencing; Chromosomal microarray analysis; Exome sequencing; Fetal structural or growth anomalies; Prenatal diagnosis

# INTRODUCTION

Congenital fetal anomalies occur in approximately 3% of pregnancies[1], and many of these anomalies have an underlying genetic etiology. Identification of the genetic basis of the anomalies enables informed decision-making, improves perinatal care, and helps to assess recurrence risk for future pregnancies. Chromosomal microarray analysis (CMA) has been broadly adopted to detect copy number variants (CNVs) in prenatal diagnoses with an additional 6% of diagnostic yield over standard karyotyping in fetuses with structural anomalies observed by ultrasound[2, 3]. CMA can detect CNVs as small as 10-100 kb in length depending on the probe density. However, smaller variants, such as single nucleotide variants (SNVs) and small insertions or deletions (INDELs), which also contribute to a substantial portion of genetic disorders remain undetectable by this approach[4-7]. Exome sequencing (ES), which detects SNVs, INDELs, and CNVs covering multiple exons, has been proven to be a powerful tool in prenatal diagnosis. In clinical practice, ES can be conducted in CMA-negative cases to further search for single-base lesions. Emerging studies have shown that ES has a detection rate of 8.5% to 10% in fetal structural abnormalities with normal karyotype and CMA results[8, 9].

Even though the combination of CMA and ES considerably increases the diagnostic yield[10, 11], it is a time-consuming stepwise approach and requires a large amount of DNA. Given the time-sensitive nature of the prenatal stage and the potential inaccessibility of adequate fetal samples, the clinical utility of this sequential scheme in prenatal diagnosis is undetermined. Genome sequencing (GS) is capable of detecting almost all types of genomic variants with a low input-DNA requirement (~100 ng) and is proposed to be beneficial in prenatal diagnosis[12, 13].

Several studies have supported GS as the first-tier test in the postnatal evaluation of patients with unexplained developmental delay, intellectual disability, congenital anomalies, and autism spectrum

disorder[14-16]. However, except for small case series with a limited number of highly selected fetuses, few studies have tried to evaluate the feasibility and value of GS in prenatal diagnosis. Talkowski et al.[17] identified precise translocation breakpoints that directly disrupted *CHD7* and *LMBRD1* by using ~12-fold GS in an undiagnosed prenatal sample, and this finding could not have been reliably inferred from conventional karyotyping. Choy et al.[18] demonstrated that 30-fold GS could provide a two-fold increase in diagnostic yield (32.0%, 16/50) in fetuses with increased nuchal translucency ( $\geq 3.5$  mm) compared with routine CMA and/or karyotyping (16.0%, 8/50). The additional diagnoses by GS include 7 (14%, 7/50) cases with pathogenic or likely pathogenic (P/LP) SNVs/INDELs, and their result demonstrated the potential of GS to replace the combination of CMA and ES in prenatal diagnosis.

To comprehensively investigate the diagnostic yield of GS and the feasibility of GS as a first-line approach in prenatal diagnosis, we prospectively applied GS to 111 fetuses with structural or growth anomalies and compared the results of GS with CMA plus ES.

## PATIENTS AND METHODS

### Study design and participants

The flowchart of this study is presented in Figure 1. The parents of 111 fetuses with structural or growth anomalies identified by sonographic examination and indicated for prenatal genetic tests were recruited from Shanghai First Maternity and Infant Hospital. Fetal samples were obtained through an invasive diagnostic procedure such as amniocentesis, chorionic villus sampling or cordocentesis. Parental peripheral blood samples were collected. Samples from fetuses were subjected to two test strategies: parent-fetus trio GS and CMA plus ES. ES was sequentially performed only in samples with negative CMA results. P/LP SNVs/INDELs were validated by Sanger sequencing, while P/LP CNVs identified by GS were cross-

validated with the results of CMA. Quantitative PCR (qPCR) was conducted for additional P/LP CNVs identified by GS. Genetic counseling was performed both before and after the prenatal genetic tests, and pregnant women were informed about the results of the tests they had taken. Pregnancy outcomes were followed up. This study protocol was approved by the Ethics Committee of Shanghai First Maternity and Infant Hospital. Written informed consent was obtained from pregnant women and their partners.

## GS

Parental and fetal samples were sequenced concurrently (fetus-parent trios or dyads testing). First, 80-200 ng of genomic DNA from each sample was sheared by the Covaris S220 Focused Ultrasonicator (Covaris, Woburn, USA). The fragmented DNA was further processed with AMPure XP beads (Life Sciences, Indianapolis, USA) to obtain 100 bp-300 bp fragments. Library construction, including end repair, A-tailing, adapter ligation, and 7 cycles of PCR amplification, was subsequently conducted. The PCR products were then heat-denatured to form single-strand DNAs, followed by circularization with DNA ligase, and the remaining linear molecule was digested with the exonuclease. After construction of the DNA nanoballs, paired-end sequencing with 100 bp at each end was carried out for each sample with a minimal read depth of 40-fold on the MGISEQ-2000 platform (MGI, Wuhan, China)[19].

Data analysis and bioinformatics processing are described in detail in Supplementary Materials and Methods. Briefly, sequencing reads were aligned to the NCBI37/hg19 assembly using the Burrows-Wheeler Aligner (BWA, version 0.7.17) with default parameters. SNVs and INDELs were called using the Genome Analysis Toolkit (GATK, version 4.0.11) and annotated by the in-house pipeline BGICG\_Anno (version 0.3.9). CNVs and structural variants (SVs) were analyzed using CNVnator (version v0.3.2) and LUMPY (version 0.3.0), respectively. The outputs of these callers were merged into a single-variant call

format (VCF) file per trio and annotated with public databases and our in-house databases. The results of the SNV/INDEL, CNV and SV analyses were integrated and reviewed for the interpretation of pathogenicity.

## **CMA**

CytoScan 750K (Affymetrix, Santa Clara, USA) was the CMA platform used in the prenatal genetic diagnosis center of Shanghai First Maternity and Infant Hospital. A total of 250 ng of DNA was required for routine prenatal CMA testing according to the manufacturer's protocols[20]. CNVs were analyzed via CHAS 2.0 software (NCBI37/hg19), and the reporting size threshold of the CNVs was set at 100 kb with a marker count of  $\geq 50$ .

## **ES**

ES was also performed by inputting 150-300 ng of genomic DNA from each sample and was prepared with the same library construction procedure as GS. Exome capture using the MGIEasy Exome Capture V4 Probe (MGI, Wuhan, China) was followed by paired-end read sequencing ( $2 \times 100$  bp read length) on the MGISEQ-2000 platform with an average depth of  $\geq 100$ -fold. Exome sequencing data analysis was performed as previously described[21].

## **Data interpretation and reporting**

CNVs detected by CMA and GS were interpreted following standards of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen)[22]. Among the large numbers of SNVs/INDELs, we prioritized candidate causative SNVs/INDELs per the following criteria: 1) absent or with a minor allele frequency  $\leq 1\%$  in the databases of ExAC (<http://exac.broadinstitute.org>) and

gnomAD (<https://gnomad.broadinstitute.org>), corresponding to the variant evidence as PM2; 2) family segregation information that was consistent with the inheritance of the variants (PS2/PM6/PM3); 3) supporting evidence from published literature (e.g., PS1/PS3/PS4/PM5); 4) null variants or CNVs that overlapped with established triplosensitive, haploinsufficient genes or genomic regions (PVS1); 5) conservation and predicted impact on coding and noncoding sequence (PP3); and 6) relevance to the fetal clinical phenotype (PP4). All the selected variants were assessed for pathogenicity based on the adapted ACMG guidelines and ClinGen sequence variant interpretation working group per updated recommendations for the ACMG criteria[23-26].

All of the candidate causative variants were reviewed by a multidisciplinary team of fetal medicine specialists, genetic counselors and geneticists. The result that was unanimously agreed upon to explain the fetal phenotype included P/LP variants consistent with the inheritance pattern and phenotype of related disorders, as well as variants of unknown significance (VUS) among Online Mendelian Inheritance in Man (OMIM) disease-causing genes that matched the fetal phenotype and was found in trans with a P/LP variant in an autosomal recessive condition, was designated positive or diagnostic; otherwise, was designated negative or undiagnostic. The pregnant women and their partners were informed of the ACMG secondary findings[27], and fetal and parental incidental findings were reported only when consented in the pretest informed consent process according to the ACMG document[28].

## **Data validation**

SNVs/INDELs/SVs were validated by Sanger sequencing. For CNV validation, qPCR was conducted for additional P/LP CNVs identified by GS. The HCMV Real-Time PCR Kit (Liferiver, Shanghai, China) was used for validation of one case that was identified by GS as suspicious for CMV infection.

## RESULTS

Between November 2019 and January 2020, 111 fetuses with structural or growth anomalies and 209 matched parental samples were eligible for inclusion in our study. Samples from a total of 320 individuals (106 fetus-parental trios, including 4 sets of twins and 5 fetus-parent dyads) were analyzed by GS, while the 111 fetal samples were also analyzed by CMA in parallel. Samples from 102 fetuses with negative CMA results were subsequently analyzed by ES. The fetuses were assessed at a median gestational age of 24 (range 16–34) weeks. The results of the Down syndrome maternal-serum screening test or noninvasive prenatal screening test were recorded when available. The detailed clinical information is available in Table S1.

### Comparison of GS and CMA plus ES

With CMA plus ES strategy, 6 (5.4%, 6/111) fetuses with aneuploidies (3 with trisomy 21, 2 with trisomy 13 and 1 with trisomy 18) and 3 (2.7%, 3/111) fetuses with P/LP CNVs were identified by CMA (Table 1), while 13 (11.7%, 13/111) fetuses with P/LP SNVs/INDELs were identified by ES (Table 2), providing a diagnostic yield of 19.8% (22/111). In comparison, GS not only detected all genetic variants in 22 cases diagnosed by CMA plus ES but also reported more types of variants and provided more comprehensive genetic information (Figure 2). Specifically, in cases 102-1 and 102-2, which were twins with severe hydroderma and pleural effusion, a 19.7-Mb deletion on the terminal q-arm of chromosome 13 (13q32.1q34), in which 3 genes were associated with autosomal dominant developmental disorder, and a 1.6-Mb duplication on the terminal q-arm of chromosome 15 (15q26.3) were both identified by CMA and GS (Figure 3A). However, GS data indicated that these two CNVs were caused by an imbalanced translocation between chromosome 13 and chromosome 15 in both twin fetuses when we checked the



precise breakpoints of the detected CNVs through Integrative Genomics Viewer (IGV). This imbalanced translocation was then confirmed by Sanger sequencing (Figure S1). The fetal karyotype strongly suggests parental translocations between chromosomes 13 and 15, and indeed, a balanced translocation was identified in the father of the twins in the analysis of the GS data, while their mother's karyotype was normal. We further delineated aberrations in the breakpoint region by using IGV and confirmed the breakpoints of the balanced interchromosomal translocation by Sanger sequencing (Figure 3B). There was a templated 13-bp duplication and a 34-bp deletion in the breakpoint junction on chromosomes 13 and 15, respectively (Figure 3C), indicating that the breakpoint repair mechanism might be caused by nonhomologous end joining with a high degree of precision but occasional deletion during DNA processing before ligation[29]. Therefore, for this family, GS provided comprehensive and precise genetic information by detecting more types of variants that might not be detectable by CMA or ES.

In addition, the potential for GS to provide a comprehensive picture of the molecular landscape for genetic interpretation was also reflected in the concurrent detection of CNVs and SNVs. For example, in case 51 with ventricular septal defect, choledochal cyst, ventriculomegaly and echogenic crystalline lens (suspected congenital cataract), 3 *de novo* pathogenic CNVs were detected by both CMA and GS. One of these CNVs, an ~11.7-Mb duplication on chromosome 8p22-8p23.2 associated with 8p23.1 duplication syndrome[30], whose common features in the reported prenatal cases included congenital heart disease and whose novel features in the reported postnatal cases included ocular anomalies, was considered to explain the fetal phenotype; thus, ES was not carried out in this case due to the positive CMA finding. However, a clinically relevant *de novo* pathogenic SNV NM\_005267.4, c.593G>A (p. Arg198Gln) in the *GJA8* gene, which was associated with Cataract 1, multiple types (OMIM: #116200), were identified by GS (Figure 2). This SNV has been reported in three children and their affected mother. All of them were diagnosed with

bilateral cataracts soon after birth, and the variant detected in their mother via segregation analysis was *de novo*[31]. The SNV has also been reported in three affected members of an Indian family with a cataract history and was absent in 400 normal controls[32]. After comprehensive evaluation by clinicians and geneticists, both the SNV and the pathogenic CNV were considered to contribute to the fetus's phenotype. However, the SNV was missed by the current CMA plus ES strategy.

To further demonstrate the capacity of GS to detect CNVs and SNVs/INDELs, we first compared all of the CNVs with sizes  $\geq 100$  kb detected between CMA and GS since the reporting threshold of CytoScan 750K we adopted for CMA was 100 kb. GS not only detected all of the CNVs with sizes  $\geq 100$  kb, which were detected by CMA, but also detected additional CNVs that CMA did not detect (Table S2). In addition, CNVs with sizes smaller than 100 kb were also detected by GS (Table S3), but none of these small CNVs were determined to be pathogenic. Our data indicated that GS was at least as powerful as CMA in the detection of CNVs with sizes  $\geq 100$  kb and had more potential to detect smaller CNVs. Second, we compared ES and GS in terms of their abilities to detect SNVs/INDELs in the coding region. Although the mean sequencing depth was higher in ES (147.1-fold [SD 24.0] vs. 57.6-fold [SD 4.2]), the coverage at 20-fold of GS was higher (97.8% [SD 1.3] vs. 96.7% [SD 1.3]), and GS detected more candidate variants in the coding region than ES (95% CI,  $p < 0.0001$ , paired t-test, Table S4), which indicated that GS is more powerful than ES for detecting exome variants.

### **Additional findings of GS**

With clinical GS, there is potential for recognition and reporting of incidental findings unrelated to the indication for ordering sequencing and even of the genetic information of associated microorganisms due to the noncapture method of GS. In case 5, a fetus with fetal growth restriction was incidentally found by

GS to be infected with cytomegalovirus (CMV) (Figure 2), while the results of CMA and ES were both negative for CMV. This case was further validated by amplification of the viral DNA in the amniotic fluid through RT-PCR (Figure S2). CMV is the leading infectious cause of newborn malformation[33]. Inspired by this case, we developed a bioinformatics pipeline for the detection of pathogens associated with adverse perinatal outcomes, such as CMV, Toxoplasma (TOX), and Herpes simplex virus 1/2 (HSV 1/2), and analyzed all of the fetal samples; however, no additional positive cases were identified. Even so, GS showed its potential to provide information on nongenetic factors that neither CMA nor ES could provide in this case.

### **Subgroup analysis**

The 111 fetuses that were eligible for GS were categorized into 10 phenotypic groups based on the types of anomalies detected by ultrasound. The molecular diagnostic rate among 10 different phenotypic groups ranged from 0%-39.1% (9/23), as shown in Table S5. The highest diagnostic rate of 39.1% (9/23) was achieved for fetuses with multisystem anomalies, followed by 30.8% (4/13) for fetuses with cardiac anomalies. We did not find any diagnostic genetic variants in 5 fetuses with abdominal anomalies or in 7 fetuses with chest anomalies. Among 9 cases diagnosed with multisystem anomalies, 8 (88.9%) cases had chromosomal disorders, and 1 (11.1%) case had a monogenic disorder.

### **Impact on pregnancy outcome**

The effect of genetic diagnosis on pregnancy outcome is shown in Table 3. Follow-up results were available for all 111 fetuses. Among those of 22 diagnosed fetuses, the parents of 21 (95.5%, 21/22) fetuses opted for termination, and the parents of the remaining fetus (4.5%, 1/22) with clubhands, facial

abnormality, omphalocele and umbilical cord cyst opted to continue the pregnancy. Among those of 89 undiagnosed fetuses, the parents of 34 (38.2%, 34/89) fetuses opted for termination, and the parents of 55 (61.8%, 55/89) fetuses opted to continue the pregnancies (Table S1). We noticed that there were no significant differences in the severities of abnormal fetal phenotypes between the diagnosed group and the undiagnosed group. Thus, in the context of an abnormal fetal phenotype, parents opted for termination of pregnancy significantly more often when they received a positive genetic diagnosis ( $p=0.00000053$ , Fisher's exact test), which indicated that the specific genetic diagnosis influenced parental decision-making.

## DISCUSSION

In this prospective study, we applied GS in parallel with CMA plus ES to 111 fetuses with a broad range of structural or growth anomalies. To our knowledge, this is the largest prospective study on the use of high-coverage trio GS for the prenatal diagnosis of fetal structural abnormalities. GS, as a single test, revealed positive genetic findings in 22 fetuses, corresponding to a diagnostic rate of 19.8% (22 in 111). The same disease-causing variants were detected by GS and CMA plus ES in 19 fetuses of these fetuses, including 6 fetuses with chromosomal aneuploidies and 13 fetuses with P/LP SNVs or INDELs. For the remaining 3 fetuses, GS provided more comprehensive and precise genetic information than CMA plus ES, revealing twin fetuses with an imbalanced translocation and one fetus with an additional pathogenic variant in the *GJA8* gene. Furthermore, GS also incidentally identified intrauterine CMV infection in a growth-restricted fetus. Taken together, these findings show that GS can detect more types of variants and even provide clinically relevant nongenetic information that might not be detectable by CMA or ES, which indicates that GS has the potential to serve as a promising first-tier approach in prenatal diagnosis.

Previous studies[4, 6, 8, 11] have demonstrated the clinical utility of CMA and ES in prenatal diagnosis, and a combination of these two approaches for each case has been warranted. In our study, the capacity of GS for the prenatal diagnosis of fetal structural anomalies was more powerful than that of CMA plus ES in the following ways. First, precise breakpoint identification and analysis through GS data could detect more types of variants. In cases 102-1 and 102-2, twins with severe hydroderma and pleural effusion, an imbalanced translocation between chromosome 13 and chromosome 15 arising from a balanced paternal translocation was identified by GS, while CMA identified only one duplication and one deletion on these two chromosomes. The precise genetic diagnosis via GS allowed an accurate assessment of the recurrence risk in future pregnancies for this family. Second, GS can avoid the omission of the disease-causing variants by the clinical CMA plus ES sequential test. For example, in case 51, a *de novo* pathogenic SNV in the *GJA8* gene was identified by GS, while ES was not carried out due to the positive CMA results. GS provided more comprehensive genetic information in this case. Finally, GS was able to provide clues regarding the presence of nongenetic factors by incidentally identifying intrauterine CMV infection in a growth-restricted fetus. The pregnant woman was determined to be CMV-IgG positive and CMV-IgM negative before pregnancy, which indicated that her preconception immunity against CMV had been established. Given that the rate of fetal transmission in nonprimary infection is low, at 0.15-2%[34, 35], the intrauterine infection with CMV was not suspected at first; however, GS was able to provide the nongenetic information on an infection that had been overlooked in this case.

In addition, GS also demonstrated its advantages of a more rapid turnaround time (TAT) and lower input-DNA requirement in this study. The diagnostic process of CMA plus ES is time-consuming, with a median TAT of 10 (SD 2) days for CMA and 21 (SD 6) days for ES, and a total amount of ~400 ng DNA would be required from limited prenatal samples. However, the GS approach requires a lower amount of

fetal DNA (100 ng) and provides a more rapid median TAT of 18 (SD 6) days (Table S6), which is meaningful for such a time-sensitive analysis[36] since it reduces anxiety and stress for families and allows couples to have adequate time for genetic counseling and decision-making.

Of note, the prenatal detection rate for P/LP CNVs (2.7%) in our study was significantly lower than those in previous large-scale studies (4.3%-8.2%)[2, 11, 37]. This may be because the majority of fetuses (80.2%, 89 of 111 fetuses) in our study received serum screening or noninvasive prenatal screening (NIPS), and common P/LP CNVs were excluded before referral to our fetal medicine unit and prenatal diagnosis center. It is important to note that 6 fetuses with chromosomal aneuploidy underwent GS in our study. We speculated that there are two reasons for this. 1) The sonographic examination determined fetal abnormalities at an earlier gestational age before the serum screening and NIPS test, and these fetuses met our inclusion criteria and were included in our study cohort. 2) In resource-limited settings or due to the personal preference of the pregnant woman, serum screening and NIPS were not offered, and their fetuses were determined to be abnormal afterward. These cases were also included. Indeed, in our cohorts, 22 fetuses referred for CMA did not undergo serum screening or NIPS, and this resulted in 6 (27.3%) fetuses with trisomy 21/18/13 still being identified during GS implementation. We also note that the prenatal detection rate for P/LP SNVs/INDELs in our study was lower than that in a previous study of 50 fetuses with increased nuchal translucency (11.7% vs 14%)[18], possibly due to our larger cohort and the broader range of fetal structural anomalies in our study. However, the incremental diagnostic yield of ES in CMA-negative cases was higher in this study (12.7% vs 8.5%-10%)[8, 9] than in previous prospective studies on the prenatal application of ES in fetuses with structural or growth anomalies who had normal CMA results. One possible reason is the smaller cohort of this study.

However, the clinical usefulness of our GS scheme may be further improved, and many

considerations should be addressed before implementing GS in clinical practice. First, we noticed that 6 fetuses with chromosomal aneuploidy went through our GS process. Thus, a reliable, rapid and cost-effective method, such as quantitative fluorescence polymerase chain reaction (QF-PCR), was needed to preliminarily exclude common fetal chromosomal aneuploidies before GS in clinical application, which could optimize the cost-effectiveness of the analysis[38]. Second, a large number of VUS variants in the noncoding region and incidental findings are called by GS, and a major concern is whether they should be reported prenatally. This decision is a particularly challenging because of the uncertainty associated with these findings, which further increases the complexity of genetic counseling compared to ES and increases difficulty in parental decision-making. Additionally, the choice to report the finding should also be weighed against the risk of missing a potential molecular diagnosis if not reported. Finally, even though sequencing costs continue to decrease, the cost of GS is comparatively high, especially for trio sequencing. Studies evaluating the cost-effectiveness of GS are warranted before clinical implementation.

In conclusion, our data show that GS provides comprehensive detection of various genomic variants in fetuses with structural or growth anomalies. In lieu of two separate analyses, GS, as a single test with a rapid TAT, should be performed as it has an equivalent diagnostic rate to that of CMA plus ES and provides a more comprehensive picture of the molecular landscape for the explanation of the mechanism. Although prospective studies with larger cohorts are warranted, our study demonstrates the feasibility of replacing CMA plus ES with GS as a first-line test for prenatal diagnosis in the near future.

## ACKNOWLEDGMENTS

We wish to extend our deepest gratitude to the individuals who participated in this research. Authors from Shanghai First Maternity and Infant Hospital designed the study, enrolled patients and collected samples,

determined the detailed fetal clinical phenotype, performed CMA detection and interpretation, provided genetic counseling, followed up patients and revised the manuscript. BGI Genomics codesigned the study, performed GS and ES detection and interpretation, analyzed and summarized data, drafted the initial manuscript, and revised the manuscript. The final diagnosis was determined by a multidisciplinary team of fetal medicine specialists, genetic counselors and geneticists from the above cooperation units. This work was supported by a grant from the National Key Research and Development Program of China (No.2018YFC1002900). This work was also supported by China National GeneBank (CNGB). The data that support the findings of this study have been deposited into CNGB Sequence Archive [39] of CNGBdb [40] with accession number CNP0001235.

## DISCLOSURE

The authors declare no conflicts of interest.

## REFERENCE

1. Murphy, S.L., et al., *Annual Summary of Vital Statistics: 2013-2014*. Pediatrics, 2017. **139**(6).
2. Wapner, R.J., et al., *Chromosomal microarray versus karyotyping for prenatal diagnosis*. N Engl J Med, 2012. **367**(23): p. 2175-84.
3. Callaway, J.L., et al., *The clinical utility of microarray technologies applied to prenatal cytogenetics in the presence of a normal conventional karyotype: a review of the literature*. Prenat Diagn, 2013. **33**(12): p. 1119-23.
4. Meier, N., et al., *Exome sequencing of fetal anomaly syndromes: novel phenotype-genotype discoveries*. Eur J Hum Genet, 2019. **27**(5): p. 730-737.
5. Vora, N.L., et al., *Prenatal exome sequencing in anomalous fetuses: new opportunities and challenges*. Genet Med, 2017. **19**(11): p. 1207-1216.
6. Drury, S., et al., *Exome sequencing for prenatal diagnosis of fetuses with sonographic abnormalities*. Prenat Diagn, 2015. **35**(10): p. 1010-7.
7. Normand, E.A., et al., *Clinical exome sequencing for fetuses with ultrasound abnormalities and a suspected Mendelian disorder*. Genome Med, 2018. **10**(1): p. 74.
8. Lord, J., et al., *Prenatal exome sequencing analysis in fetal structural anomalies detected by*



- ultrasonography (PAGE): a cohort study*. Lancet, 2019. **393**(10173): p. 747-757.
9. Petrovski, S., et al., *Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study*. Lancet, 2019. **393**(10173): p. 758-767.
10. Dharmadhikari, A.V., et al., *Copy number variant and runs of homozygosity detection by microarrays enabled more precise molecular diagnoses in 11,020 clinical exome cases*. Genome Med, 2019. **11**(1): p. 30.
11. Fu, F., et al., *Whole exome sequencing as a diagnostic adjunct to clinical testing in fetuses with structural abnormalities*. Ultrasound Obstet Gynecol, 2018. **51**(4): p. 493-502.
12. Saunders, C.J., et al., *Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units*. Sci Transl Med, 2012. **4**(154): p. 154ra135.
13. Willig, L.K., et al., *Whole-genome sequencing for identification of Mendelian disorders in critically ill infants: a retrospective analysis of diagnostic and clinical findings*. Lancet Respir Med, 2015. **3**(5): p. 377-87.
14. Gilissen, C., et al., *Genome sequencing identifies major causes of severe intellectual disability*. Nature, 2014. **511**(7509): p. 344-7.
15. Lionel, A.C., et al., *Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test*. Genet Med, 2018. **20**(4): p. 435-443.
16. Lindstrand, A., et al., *From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability*. Genome Med, 2019. **11**(1): p. 68.
17. Talkowski, M.E., et al., *Clinical diagnosis by whole-genome sequencing of a prenatal sample*. N Engl J Med, 2012. **367**(23): p. 2226-32.
18. Choy, K.W., et al., *Prenatal Diagnosis of Fetuses With Increased Nuchal Translucency by Genome Sequencing Analysis*. Front Genet, 2019. **10**: p. 761.
19. Huang, J., et al., *A reference human genome dataset of the BGISEQ-500 sequencer*. Gigascience, 2017. **6**(5): p. 1-9.
20. Lin, S.B., et al., *Improved assay performance of single nucleotide polymorphism array over conventional karyotyping in analyzing products of conception*. J Chin Med Assoc, 2015. **78**(7): p. 408-13.
21. Bergant, G., et al., *Comprehensive use of extended exome analysis improves diagnostic yield in rare disease: a retrospective survey in 1,059 cases*. Genet Med, 2018. **20**(3): p. 303-312.
22. Riggs, E.R., et al., *Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen)*. Genet Med, 2020. **22**(2): p. 245-257.
23. Richards, S., et al., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology*. Genet Med, 2015. **17**(5): p. 405-24.
24. Abou Tayoun, A.N., et al., *Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion*. Hum Mutat, 2018. **39**(11): p. 1517-1524.
25. Brandt, T., et al., *Adapting ACMG/AMP sequence variant classification guidelines for single-gene copy number variants*. Genet Med, 2019.
26. Riggs, E.R., et al., *Technical standards for the interpretation and reporting of constitutional copy-*

- number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen)*. Genet Med, 2019.
27. Kalia, S.S., et al., *Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics*. Genet Med, 2017. **19**(2): p. 249-255.
28. Monaghan, K.G., et al., *The use of fetal exome sequencing in prenatal diagnosis: a points to consider document of the American College of Medical Genetics and Genomics (ACMG)*. Genet Med, 2020.
29. Carvalho, C.M. and J.R. Lupski, *Mechanisms underlying structural variant formation in genomic disorders*. Nat Rev Genet, 2016. **17**(4): p. 224-38.
30. Barber, J.C., et al., *8p23.1 duplication syndrome; common, confirmed, and novel features in six further patients*. Am J Med Genet A, 2013. **161A**(3): p. 487-500.
31. Prokudin, I., et al., *Exome sequencing in developmental eye disease leads to identification of causal variants in GJA8, CRYGC, PAX6 and CYP11B1*. Eur J Hum Genet, 2014. **22**(7): p. 907-15.
32. Devi, R.R. and P. Vijayalakshmi, *Novel mutations in GJA8 associated with autosomal dominant congenital cataract and microcornea*. Mol Vis, 2006. **12**: p. 190-5.
33. Manicklal, S., et al., *The "silent" global burden of congenital cytomegalovirus*. Clin Microbiol Rev, 2013. **26**(1): p. 86-102.
34. Fowler, K.B., et al., *The outcome of congenital cytomegalovirus infection in relation to maternal antibody status*. N Engl J Med, 1992. **326**(10): p. 663-7.
35. Stagno, S., et al., *Immunoglobulin M antibodies detected by enzyme-linked immunosorbent assay and radioimmunoassay in the diagnosis of cytomegalovirus infections in pregnant women and newborn infants*. J Clin Microbiol, 1985. **21**(6): p. 930-5.
36. Wang, H., et al., *Clinical utility of 24-h rapid trio-exome sequencing for critically ill infants*. NPJ Genom Med, 2020. **5**: p. 20.
37. Wang, J.C., et al., *Effects of platforms, size filter cutoffs, and targeted regions of cytogenomic microarray on detection of copy number variants and uniparental disomy in prenatal diagnosis: Results from 5026 pregnancies*. Prenat Diagn, 2019. **39**(3): p. 137-156.
38. Corsten-Janssen, N., et al., *A prospective study on rapid exome sequencing as a diagnostic test for multiple congenital anomalies on fetal ultrasound*. Prenat Diagn, 2020.
39. Guo, X., et al., *CNSA: a data repository for archiving omics data*. Database (Oxford), 2020. **2020**.
40. *CNGBdb: China National GeneBank DataBase, Hereditas*, DOI: 10.16288/j.ycz.20-080.

## FIGURE LEGENDS

**Figure 1.** Flowchart of the study. A total of 111 fetuses with structural or growth anomalies were subjected to two test strategies: trio genome sequencing (GS) and chromosomal microarray (CMA) plus exome sequencing (ES). ES was sequentially performed in only 102 fetuses with negative CMA results. The positive/negative rates are provided in each box.

**Figure 2.** Architecture of a mixed cohort referred for prenatal diagnosis. Each slice of the pie chart represents one individual in the prospective cases analyzed by genome sequencing (GS) and chromosomal microarray (CMA) plus exome sequencing (ES) where clinically relevant findings were identified. Types of variant are indicated by colors (aneuploidy, light blue; duplication and heterozygous (het) single nucleotide variant (SNV), blue; imbalanced translocation, red; intrauterine infection, yellow; compound heterozygous SNV, light gray; heterozygous SNV, dark gray). Additional genetic findings by GS are indicated by \* and additional nongenetic findings by GS are indicated by #.

**Figure 3.** Genome sequencing (GS) identified pathogenic copy number variants (CNVs) in twin fetuses arising from their father's balanced translocation. (A) Distributions of copy number across chromosome 13 and chromosome 15 are shown at the top and bottom, respectively. Dots in red indicate a heterozygous deletion of 19.7 Mb on chromosome 13, and dots in blue indicate a duplication of 1.6 Mb on chromosome 15 in both case 102-1 and case 102-2. The deletion affected 3 genes that are associated with autosomal dominant developmental disorder. (B) The father's karyotype of chromosome 13 and chromosome 15 are shown at the top and bottom. The distribution of base pair sequencing depth and Sanger sequencing results across the breakpoints on chromosome 13 and chromosome 15 are displayed, respectively, in the middle. There is a templated 13-bp duplication and a 34-bp deletion in the breakpoint junction on chromosomes 13

and 15. (C) A schematic drawing of the formation of balanced translocations between chromosomes 13 and 15 in the father.

**Table 1.** Summary of numerical disorder and pathogenic or likely pathogenic copy number variants (CNVs)

detected by CMA and GS.

Case ID	Clinical Indication	Serum	NIPS	CMA	GS	Added information by GS
			screening			
2	Cystic hygroma; Severe tetralogy of Fallot	N	N	arr(21)×3	seq(21)×3	
40	Ventricular septal defect; Pulmonary Atresia	N	N	arr(21)×3	seq(21)×3	
42	Ventricular septal defect	N	N	arr(13)×3	seq(13)×3	
51	Ventricular septal defect; Choledochol cyst; Ventriculomegaly; Echogenic crystalline lens	N	Low-risk	arr[GRCh37]6q25.3q27(158678581-170914297)×3 arr[GRCh37] 8p23.3(158048-2193914)×1 arr[GRCh37]8p23.2p22(2347604-14015208)×3	seq[GRCh37]dup(6)(q25.3q27)dn chr6:g.158661896-170927571dup seq[GRCh37]del(8)(p23.3)dn chr8:g.156079-2185433del seq[GRCh37]dup(8)(p23.2p22)dn chr8:g.2340301-13996893dup	A pathogenic SNV (NM_005267.4,c.593G>A,p.Arg198Gln) in <i>GJA8</i> was identified by GS
52	Congenital heart disease; Clubfeet; Atrial septal defects; tetralogy of Fallot	N	N	arr(21)×3	seq(21)×3	
84	Holoprosencephaly; Ventricular septal defect; Polydactyly	N	N	arr(13)×3	seq(13)×3	
93	Single umbilical artery;	N	N	arr(18)×3	seq(18)×3	

	Omphalocele; Umbilical					
	cord cyst; Clubhands;					
	Facial abnormality					
<b>102-1</b>	Skeletal dysplasia;	N	Low-	arr[GRCh37]13q32.1q34(95,44308	seq[GRCh37]del(13)(q32.1q34)	Arising from paternal
	Unilateral pleural effusion;		risk	8-115107733)*1	chr13:g.95440312-115169878del	balanced translocation
	Intestinal dilatation			arr[GRCh37]15q26.3(100865196-	seq[GRCh37]dup(15)(q26.3)	
				102429040)*3	chr15:g.100877325-102531392dup	
<b>102-2</b>	Fetal hydrops; Ascites;	N	Low-	arr[GRCh37]13q32.1q34(95,44308	seq[GRCh37]del(13)(q32.1q34)	Arising from paternal
	Pleural effusion;		risk	8-115107733)*1	chr13:g.95440312-115169878del	balanced translocation
	Hydroderma			arr[GRCh37]15q26.3(100865196-	seq[GRCh37]dup(15)(q26.3)	
				102429040)*3	chr15:g.100877325-102531392dup	

NIPS, Non-invasive prenatal screening; CMA, Chromosomal microarray; GS, Genome sequencing; N, No screening.

**Table 2.** Summary of pathogenic or likely pathogenic single nucleotide variants (SNVs) and small insertions or deletions (INDELs) detected by ES and GS.

Case ID	Clinical Indication	Genes	Variants (all heterozygous)	Inheritance	OMIM diseases	Novel or previously reported (PMID)
1	Oligohydramnios; Renal	<i>PKHD1</i>	NM_138694.3		#263200	
	cyst		c.7994T>C,p.Leu2665Pro	Mat	Polycystic Kidney Disease 4 with or without Polycystic Liver Disease (AR)	Reported (PMID30507656;PMID28851938)
			c.5428G>T,p.Glu1810*	Pat		Novel
3	Skeletal dysplasia;	<i>COL1A1</i>	NM_000088.3	De novo	#166200	Novel
	Micromelia		c.4280_4283delTTGA,p.Ile1427Asnfs*98		Osteogenesis Imperfecta (AD)	
10	Fetal growth restriction;	<i>SBDS</i>	NM_016038.2		#260400	Both Reported
	Cerebellar dysplasia		c.183_184delinsCT,p.Lys62*	Pat	Shwachman-Diamond Syndrome (AR)	(PMID15769891;PMID18478597;P
			c.258+2T>C	Mat		MID15284109)
12	Bilateral pleural	<i>MUSK</i>	NM_005592.3		#208150	
	effusion; Hydroderma;		c.790C>T,p.Arg264*	Mat	Akinesia Deformation	Novel
	Anhydramnios		c.1003_1006delGTTT,p.Val335Phefs*24	Pat	Sequence (AR)	Novel
15	Kidney agenesis	<i>HNF1B</i>	NM_000458.2	De novo	#137920	Reported
			c.494G>A,p.Arg165His		Renal Cysts and Diabetes Syndrome	(PMID24254850;PMID27838256;P
						MID22051731)

					(AD)	
<b>16</b>	Rhabdomyomas;	<i>TSC2</i>	NM_000548.3	De novo	#613254	Reported
	subependymal nodules		c.4258_4261delTCAG,p.Ser		Tuberous sclerosis-2	(PMID26252095;PMID10533067;P
			1420Glyfs*55		(AD)	MID29740858)
<b>30-1</b>	Rhabdomyomas;	<i>TSC2</i>	NM_000548.3	De novo	#613254	Reported (PMID27494029)
	Intracardiac echogenic		c.4762C>T,p.Gln1588*		Tuberous sclerosis-2	
	focus				(AD)	
<b>30-2</b>	Rhabdomyomas;	<i>TSC2</i>	NM_000548.3	De novo	#613254	Reported (PMID27494029)
	Intracardiac echogenic		c.4762C>T,p.Gln1588*		Tuberous sclerosis-2	
	focus				(AD)	
<b>51<sup>#</sup></b>	Ventricular septal defect;	<i>GJA8</i>	NM_005267.4	De novo	#116200	Reported (PMID16604058)
	Choledochol cyst;		c.593G>A,p.Arg198Gln		Cataract 1, multiple	
	Ventriculomegaly;				types (AD)	
	Echogenic crystalline					
	lens					
<b>64</b>	Anhydramnios;	<i>RYR1</i>	NM_000540.2		#255320	
	Hydroderma		c.6082C>T,p.Arg2028*	Pat	Minicore myopathy with	Novel
			c.165+5G>A	Mat	external	Novel
					ophthalmoplegia (AR)	
<b>77</b>	Cleft lip/palate;	<i>CHD7</i>	NM_017780.3	De novo	#214800	Novel
	Persistent left umbilical		c.2881delG,p.Glu961Serfs*1		CHARGE syndrome	
	vein		6		(AD)	



<b>97</b>	Holoprosencephaly;	<i>ZIC2</i>	NM_007129.3	De novo	#609637	Novel
	Agenesis of the corpus callosum		c.916G>T,p.Glu306*		Holoprosencephaly (AD)	5
<b>104</b>	Fetal growth restriction	<i>IARSI</i>	NM_013417.2		#617093	
			c.2420C>G,p.Pro807Arg	Mat	Growth retardation,	Novel
			c.2975A>G,p.Asn992Ser	Pat	impaired intellectual development, hypotonia, and hepatopathy (AR)	Novel
<b>112</b>	Cleft lip/palate; Small stomach; Ventricular septal defect	<i>CHD7</i>	NM_017780.3	De novo	#214800	Novel
			c.7153C>T,p.Gln2385*		CHARGE syndrome (AD)	

Mat, maternal; Pat, paternal; AR, Autosomal recessive; AD, Autosomal dominant; SNV, single nucleotide variant; ES, Exome sequencing; GS, Genome sequencing.

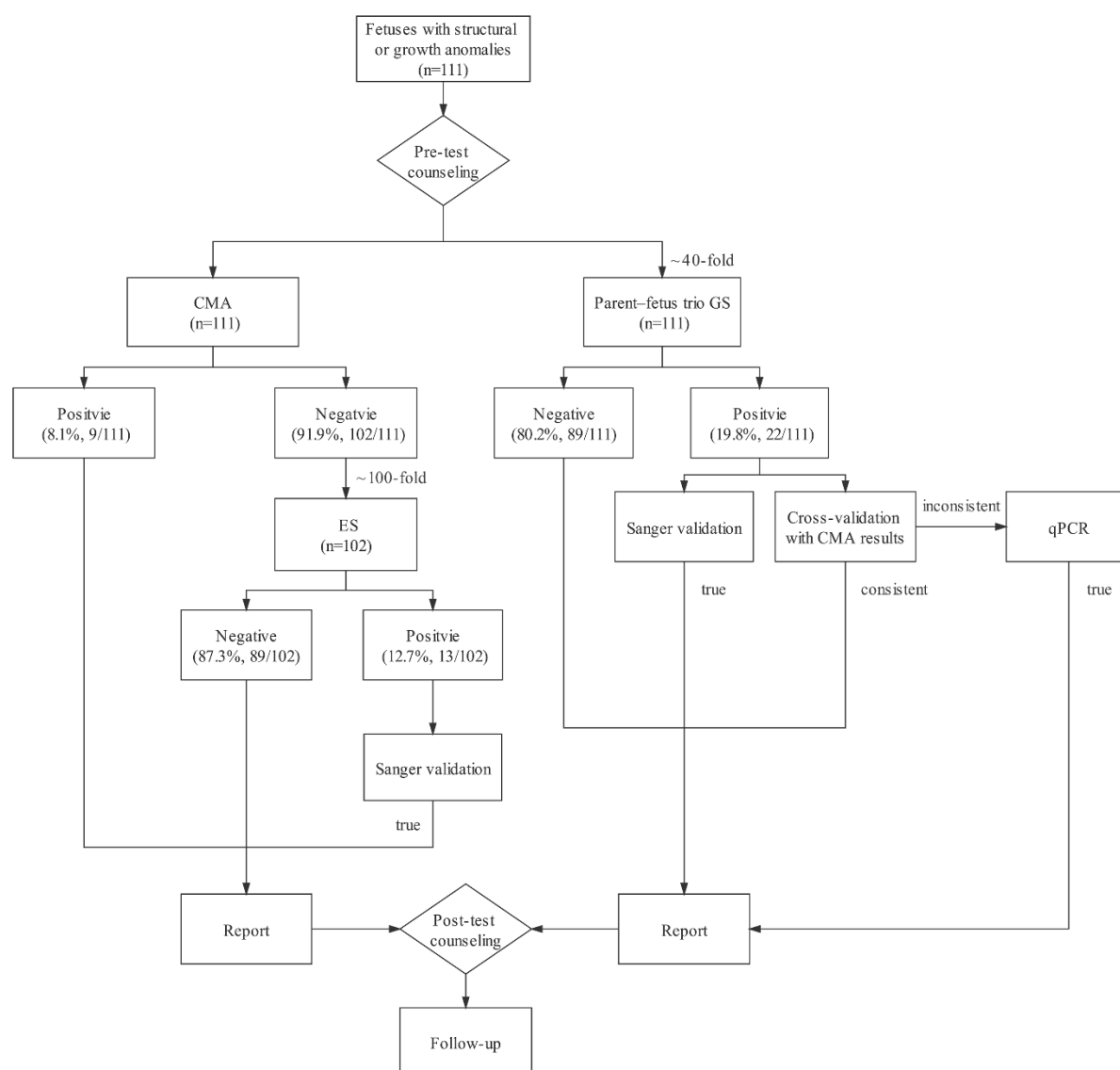
# GS identified a clinically relevant pathogenic SNV in *GJA8* gene in this CMA-positive case.

**Table 3.** Impact of genetic diagnoses on pregnancy outcome.

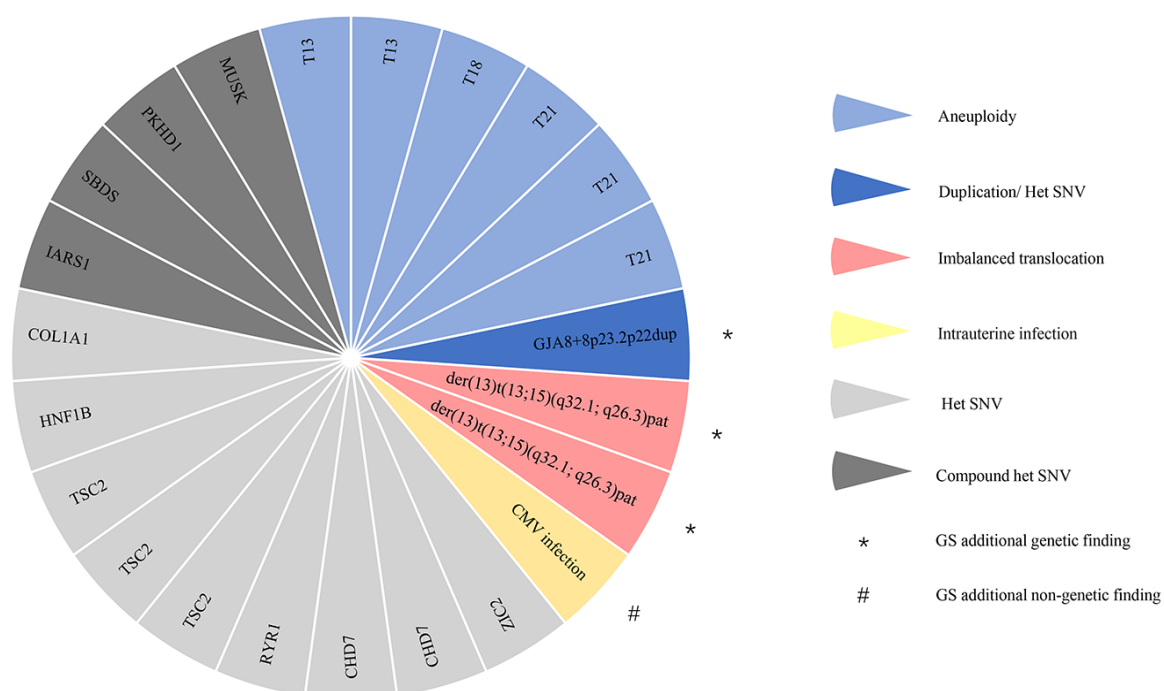
Groups	Number of cases	Termination of pregnancy		Continuing pregnancy		P value <sup>^</sup>
		Number of cases	95% CI <sup>#</sup>	Number of cases	95% CI <sup>#</sup>	
Diagnosed	22 (19.8%)	21 (95.5%)	77.2%-99.9%	1 (4.5%)	0.1%-22.8%	0.00000053
Undiagnosed	89 (80.2%)	34 (38.2%)	28.1%-49.1%	55 (61.8%)	50.9%-71.9%	
Overall	111	55 (49.6%)	39.9%-59.2%	56 (50.4%)	40.8%-60.1%	-

<sup>#</sup>95% confidence interval was calculated by binomial exact calculation.

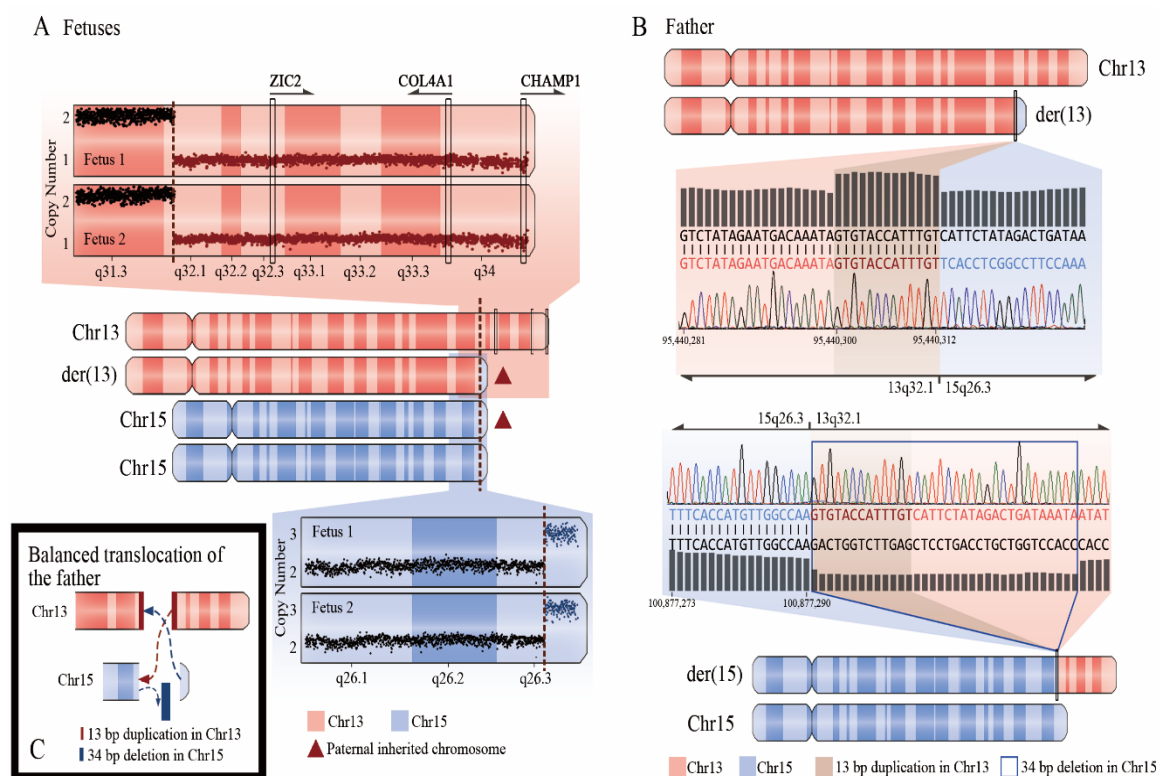
<sup>^</sup>Fisher's exact test.



**Figure 1.** Flowchart of the study. A total of 111 fetuses with structural or growth anomalies were subjected to two test strategies: trio genome sequencing (GS) and chromosomal microarray (CMA) plus exome sequencing (ES). ES was sequentially performed in only 102 fetuses with negative CMA results. The positive/negative rates are provided in each box.



**Figure 2.** Architecture of a mixed cohort referred for prenatal diagnosis. Each slice of the pie chart represents one individual in the prospective cases analyzed by genome sequencing (GS) and chromosomal microarray (CMA) plus exome sequencing (ES) where clinically relevant findings were identified. Types of variant are indicated by colors (aneuploidy, light blue; duplication and heterozygous (het) single nucleotide variant (SNV), blue; imbalanced translocation, red; intrauterine infection, yellow; compound heterozygous SNV, light gray; heterozygous SNV, dark gray). Additional genetic findings by GS are indicated by \* and additional nongenetic findings by GS are indicated by #.



**Figure 3.** Genome sequencing (GS) identified pathogenic copy number variants (CNVs) in twin fetuses arising from their father's balanced translocation. (A) Distributions of copy number across chromosome 13 and chromosome 15 are shown at the top and bottom, respectively. Dots in red indicate a heterozygous deletion of 19.7 Mb on chromosome 13, and dots in blue indicate a duplication of 1.6 Mb on chromosome 15 in both case 102-1 and case 102-2. The deletion affected 3 genes that are associated with autosomal dominant developmental disorder. (B) The father's karyotype of chromosome 13 and chromosome 15 are shown at the top and bottom. The distribution of base pair sequencing depth and Sanger sequencing results across the breakpoints on chromosome 13 and chromosome 15 are displayed, respectively, in the middle. There is a templated 13-bp duplication and a 34-bp deletion in the breakpoint junction on chromosomes 13 and 15. (C) A schematic drawing of the formation of balanced translocations between chromosomes 13 and 15 in the father.