nature medicine

LETTERS

https://doi.org/10.1038/s41591-018-0334-x

Non-invasive prenatal sequencing for multiple Mendelian monogenic disorders using circulating cell-free fetal DNA

Jinglan Zhang[1,10*], Jianli Li[2,10], Jennifer B. Saucier[3], Yanming Feng[2], Yanjun Jiang[2], Jefferson Sinson[2], Anne K. McCombs[4], Eric S. Schmitt[2], Sandra Peacock[2], Stella Chen[2], Hongzheng Dai[1], Xiaoyan Ge[1], Guoli Wang[2], Chad A. Shaw[1,5], Hui Mei[1], Amy Breman[1], Fan Xia[1], Yaping Yang[1], Anne Purgason[2], Alan Pourpak[2], Zhao Chen[1], Xia Wang[1], Yue Wang[1], Shashikant Kulkarni[1], Kwong Wai Choy[6,7], Ronald J. Wapner[8], Ignatia B. Van den Veyver[1,9], Arthur Beaudet[1], Sheetal Parmar[3], Lee-Jun Wong[1] and Christine M. Eng[1]

Current non-invasive prenatal screening is targeted toward the detection of chromosomal abnormalities in the fetus[1,2]. However, screening for many dominant monogenic disorders associated with de novo mutations is not available, despite their relatively high incidence[3]. Here we report on the development and validation of, and early clinical experience with, a new approach for non-invasive prenatal sequencing for a panel of causative genes for frequent dominant monogenic diseases. Cell-free DNA (cfDNA) extracted from maternal plasma was barcoded, enriched, and then analyzed by next-generation sequencing (NGS) for targeted regions. Low-level fetal variants were identified by a statistical analysis adjusted for NGS read count and fetal fraction. Pathogenic or likely pathogenic variants were confirmed by a secondary amplicon-based test on cfDNA. Clinical tests were performed on 422 pregnancies with or without abnormal ultrasound findings or family history. Follow-up studies on cases with available outcome results confirmed 20 true-positive, 127 true-negative, zero false-positive, and zero-false negative results. The initial clinical study demonstrated that this non-invasive test can provide valuable molecular information for the detection of a wide spectrum of dominant monogenic diseases, complementing current screening for aneuploidies or carrier screening for recessive disorders.

Since the discovery of circulating cell-free fetal DNA in the late 1990s, non-invasive prenatal screening (NIPS) for the detection of common fetal chromosomal aneuploidies using a simple blood draw from a pregnant woman has been developed[4–7]. NIPS has much higher test sensitivity and specificity than traditional maternal serum screening for trisomy 21, 18 and 13 (refs. [1,2,8]). NIPS is also increasingly used for screening for sex chromosome aneuploidies and microdeletions[9–12]. In addition, population-based carrier screening has been implemented since the 1970s, with demonstrated cost-effectiveness for reducing severe recessive monogenic diseases[13,14]. We have reported that nearly 60% of severe postnatal monogenic diseases are dominant disorders, with the majority of them caused by de novo mutations[15], but population-based prenatal screening for such diseases has not been made available despite their relative high prevalence[3]. Notably, next-generation sequencing (NGS) provides an accurate and flexible approach to non-invasive prenatal diagnosis of common dominant skeletal dysplasias[16]. Targeted non-invasive prenatal testing has been developed for monogenic disorders by using digital PCR, Sanger sequencing or NGS for a small number of variants or genes[16–18]. These methods are limited by their focus on mutations in small targeted regions and the use of locus-specific primers, which prevent concurrent detection of many sporadic mutations in highly fragmented cfDNA.

We designed a new NIPS approach for the detection of de novo or paternally inherited disease-causing variants in 30 genes associated with frequent human dominant monogenic disorders that can lead to significant adverse health outcomes at an accumulative prevalence of ~1 in 600 (Supplementary Table 1). During the library construction and sequencing processes for the NGS analysis, low-level sequence changes can be introduced that lead to sequence background noise[19–22]. To reduce such sequence noise, we used a unique molecular indexing (UMI, or molecular barcoding) technique for the NGS analysis (Fig. 1a,b). After hybridization-based target enrichment, adequate sequencing was achieved such that most UMI-labeled DNA molecules had at least two duplicated reads to assess the utility of UMI (Fig. 1c). Errors introduced during NGS library preparation and sequencing steps were significantly reduced by the UMI method for variants called with variant allele fractions $\leq 2.5\%$ ($P < 0.05$, Fig. 1d). Next, we developed a single nucleotide polymorphism (SNP)-based fetal fraction (FF) calculation method using informative transmitted parental alleles in the fetal cfDNA present in maternal plasma (see Methods and Supplementary Table 2). We compared

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. [2]Baylor Genetics, Houston, TX, USA. [3]Natera, Inc., San Carlos, CA, USA. [4]Office of Clinical Research, Baylor College of Medicine, Houston, TX, USA. [5]Department of Statistics, Rice University, Houston, TX, USA. [6]Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Hong Kong, China. [7]The Chinese University of Hong Kong-Baylor College of Medicine Joint Center For Medical Genetics, Hong Kong, China. [8]Department of Obstetrics and Gynecology, Columbia University, New York, NY, USA. [9]Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, TX, USA. [10]These authors contributed equally: J. Zhang, J. Li. *e-mail: jinglanz@bcm.edu

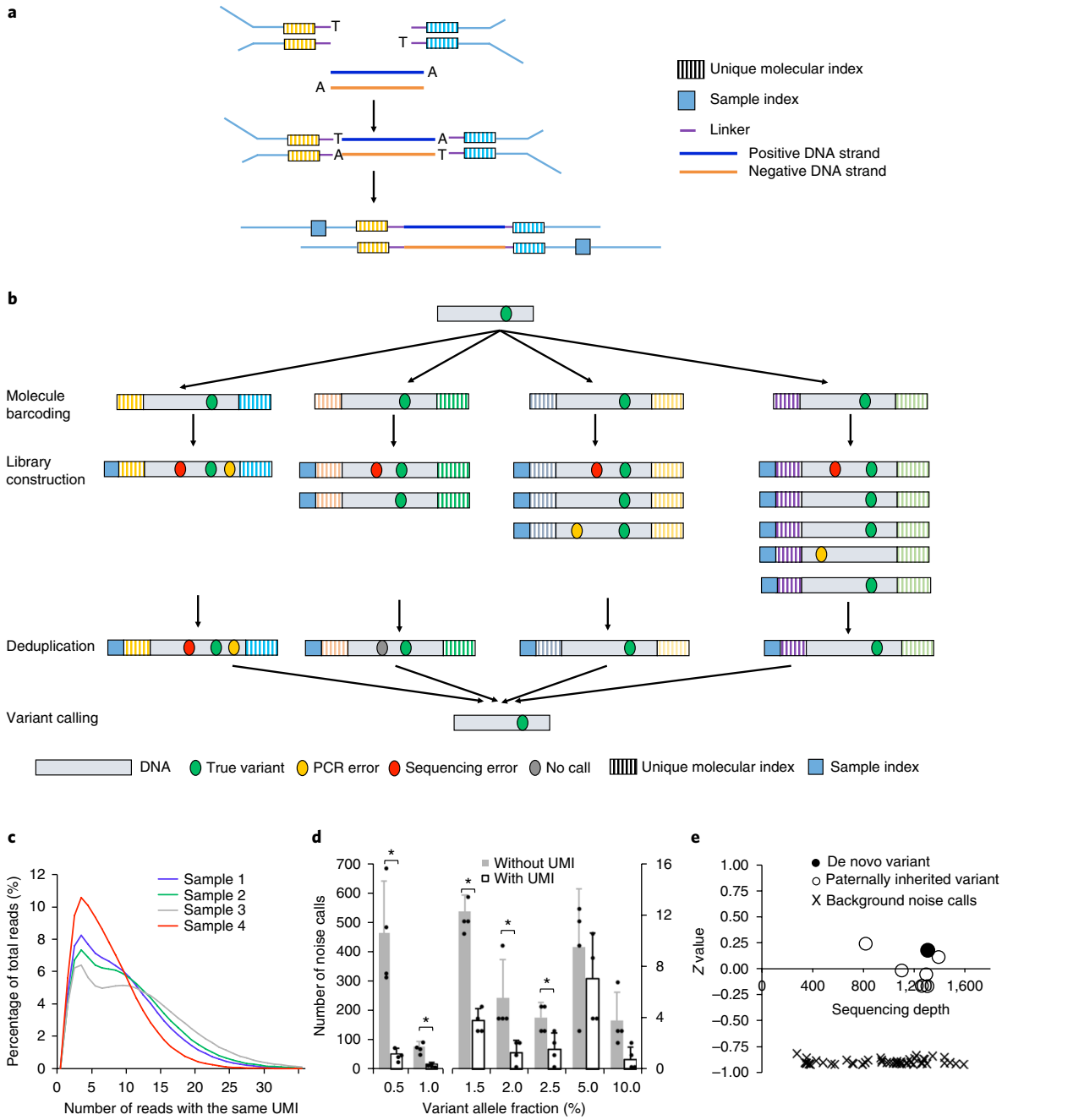


Fig. 1 | Development of non-invasive prenatal sequencing by applying UMI. a, Design of UMI for the test. Single DNA molecules were first tagged through ligation with custom-designed Y-shaped adapters with a UMI which consists of 6-bp degenerate nucleotides as molecular barcode or unique molecular index. The individual samples were then indexed using Illumina primers containing 8-bp sample indexes by PCR. **b**, Diagram depicting molecular indexing and library construction for error suppression. A deduplication algorithm was used after sequencing for the consolidation of NGS reads sharing the same UMI. Variant calling was then performed based on consensus reads after deduplication. **c**, Distribution of duplicated NGS reads with the same UMIs in four distinct representative samples used to test when the NGS reads had high duplication rate. Adequate sequencing depth was achieved such that most UMI-labeled DNA molecules had at least two duplicated reads. Experiments were performed in three independent runs demonstrating similar results. **d**, The number of background noise variant calls (analytical false variant calls) was reduced by the application of UMIs. The means in each group with different VAFs is shown in bar graphs with error bars representing s.d. ($n=4$). Statistical significance was determined by an unpaired Student's t test. Data were considered significant when the two-tailed P values were <0.05 (95% confidence interval, df = 6) as shown in those variants (labeled with *) called with VAF ≤ 2.5% ($P_{VAF0.5}=0.004$, $P_{VAF1.0}=0.0007$, $P_{VAF1.5}=0.0001$, $P_{VAF2.0}=0.04$, $P_{VAF2.5}=0.03$, $P_{VAF5.0}=0.4$, $P_{VAF10.0}=0.05$). **e**, Z-value plot for de novo, paternally inherited and false analytical variants in a representative validation sample. The value on the x-axis is the sequencing depth (total counts of the UMIs) at a locus where a variant is called.

the performance of the SNP-based FF calculation with the results obtained from a Y-chromosome marker method on 31 samples collected from pregnancies with male fetuses ($R=0.981$, Extended Data Fig. 1a) and spike-in studies using mixed DNA of proband-mother pairs with the proband DNA fractions ranging from 1.0 to 20.0% ($R=0.980$, Extended Data Fig. 1b). In addition, we compared the

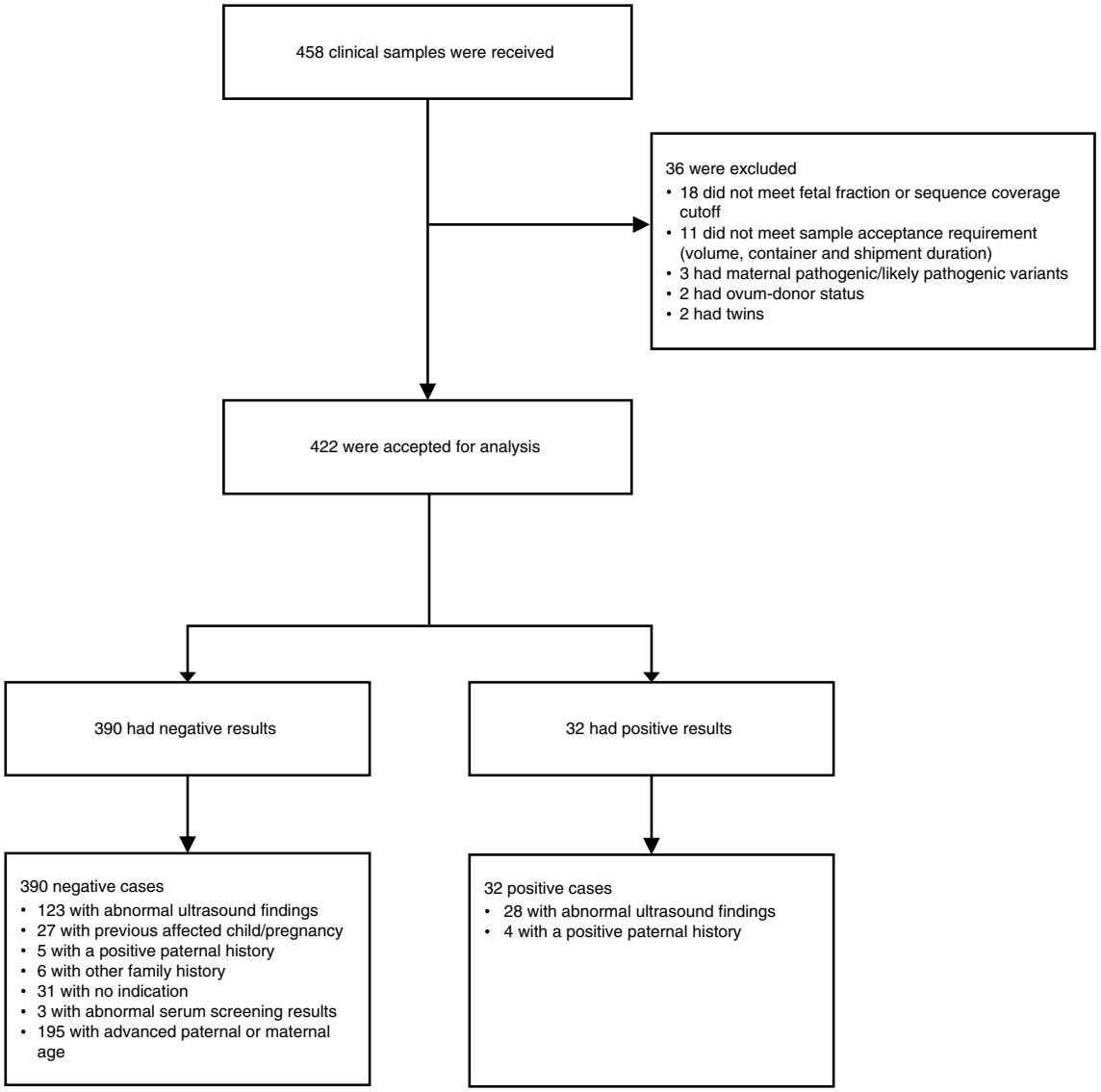


Fig. 2 | STARD (standards for reporting of diagnostic accuracy studies) flow diagram for non-invasive prenatal sequencing for multiple Mendelian monogenic disorders. A total of 458 clinical samples were received from 131 clinics in the United States, Europe, and Asia in this study. Of these samples, 422 met the acceptance criteria to complete the analysis. The sample rejection criteria as well as a summary for the negative or positive results are shown.

calculated FF for 67 cfDNA samples collected from pregnant women (FF ranging 5.0–27.6%) with the results analyzed by an independent method[23,] which yielded concordant results ($R=0.972$, Extended Data Fig. 1c). Finally, we applied a standard normal transformation process to convert the number of UMI-based NGS reads containing a locus-specific DNA variant to a Z-value (see Methods). Using this method, a Z-value cutoff ($Z>-0.6$) was set to filter out false-positive calls based on an analysis of 43 cfDNA samples in which all paternally inherited variants and confirmed de novo variants deemed true positives had a Z-value >-0.6 (Fig. 1e and Extended Data Fig. 2). To evaluate analytical sensitivity, we searched the 30 genes included in the sequencing panel for de novo or obligate heterozygous fetal variants in 76 validation samples. Those apparently de novo variants identified in the primary assay were subjected to a secondary targeted assay using the cfDNA (Supplementary Table 3). An added benefit of having a secondary assay concurrently performed for both the maternal leukocyte and plasma cfDNA is to help examine whether there is low-level maternal mosaicism, which is a potential cause of false-positive fetal results in the primary assay. False-negative calls were defined as variants (either paternally inherited or de novo variants) that should have been present but were not detected in the cfDNA. There were no analytical false-positive or negative results in the 76 samples used for test validation (Supplementary Table 4). Intra-assay, inter-assay and inter-person reproducibility studies for five samples were also performed, all of which produced consistent results (Supplementary Table 5).

A total of 458 clinical samples were received from 131 clinics in the United States, Europe, and Asia. Of these samples, 422 met the acceptance criteria (see Methods and Fig. 2). The average gestational age at the time of collection was 16.8 weeks (ranging 9.0–38.3 weeks, Fig. 2 and Supplementary Table 6). There were 241 (57.1%) women of advanced maternal age (≥35 years) and 201 men (47.6%) of advanced paternal age (≥40 years). Of the 422 total women tested, 151 (35.8%) had reported abnormal prenatal ultrasound findings indicative of a fetal developmental abnormality, 3 (0.7%) had abnormal serum screening results, and 43 (10.2%) reported a positive family history for genetic disease (Supplementary Table 6). Among 151 cases with an abnormal prenatal ultrasound, 28 yielded

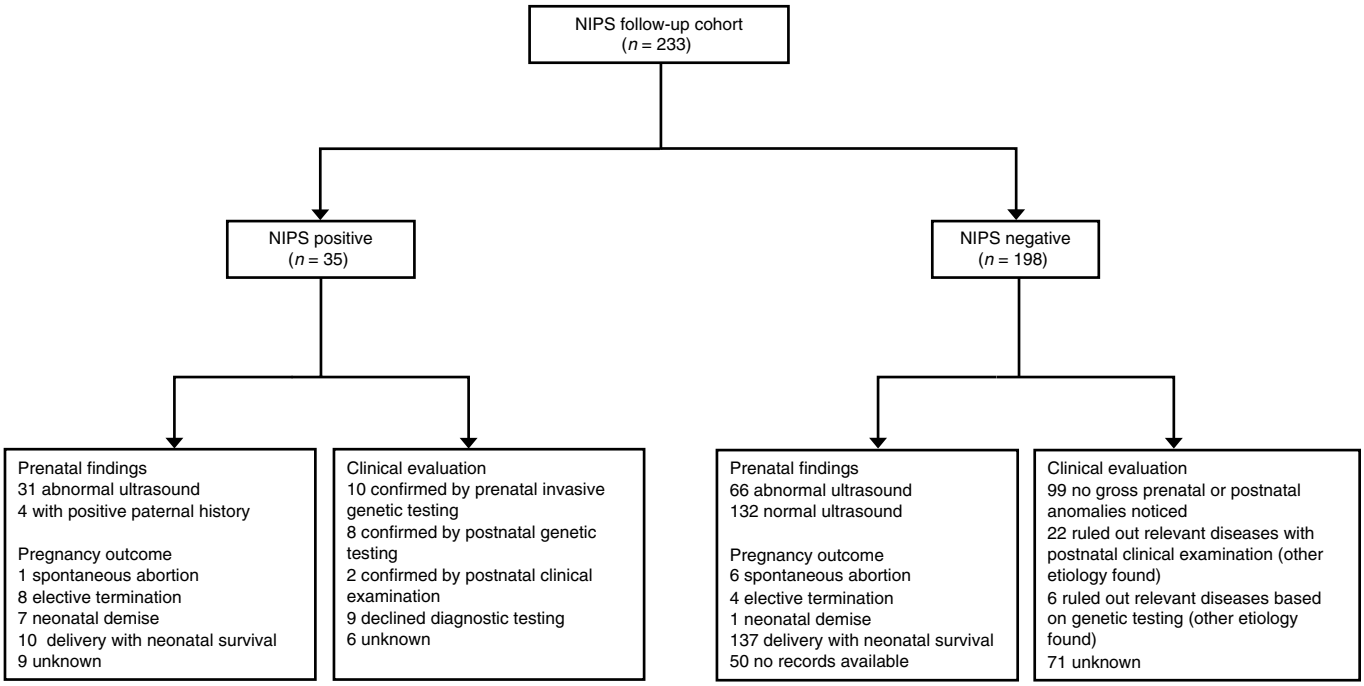


Fig. 3 | The outcome study in the NIPS follow-up cohort. Providers participated in the follow-up study for a total of 233 cases comprising 76 validation and 157 clinical cases. The positive cases in the follow-up cohort included 3 cases collected for test validation and 32 in clinical studies. The negative cases included 73 cases collected for validation and 125 cases in clinical studies. For the remaining 71 negative cases, diagnostic data were not available, largely because these pregnancies were not managed by the ordering physician at the time of delivery. The underlined cases were used for the clinical sensitivity and specificity calculation.

a positive result for a de novo pathogenic or likely pathogenic variant identified in one of the 30 genes on the screening panel (Table 1 and Supplementary Table 7). In these positive cases, the mean maternal and paternal ages were 31.7 and 35.6 years, respectively. The yield of positive molecular findings was highest for fetuses with skeletal abnormalities (19 of 59 cases, 32.2%), which included limb, digit, skull, and trunk anomalies (Supplementary Table 7). The other phenotypic categories with positive molecular findings included increased nuchal translucency or cystic hygroma (2 of 24 cases, 8.3%), cardiac defects (2 of 11 cases, 18.2%), craniofacial abnormality (2 of 8 cases, 25.0%), and multisystem anomalies (3 of 12 cases, 25.0%, Supplementary Table 7). Notably, the de novo variants identified in these 28 cases were all consistent with diseases related to the fetal phenotype seen on prenatal sonograms (Table 1). For example, a de novo pathogenic variant in the *COL1A1* or *COL1A2* gene was detected in subjects P1–P8, who all carried a fetus with skeletal abnormalities including short, bowed, or broken bones. In subjects P9–P18, de novo pathogenic variants in the *FGFR3* gene associated with achondroplasia and thanatophoric dysplasia were identified, consistent with fetal sonograms demonstrating short femur, frontal bossing, or small chest-to-abdominal circumference and femur length-to-abdominal circumference ratios. In P19 and P20, de novo pathogenic variants were found in the *FGFR2* gene associated with craniosynostosis, consistent with abnormal fetal skull shape. In P22, a de novo pathogenic variant, c.1435C>T (p.R479*) in the *NIPBL* gene associated with Cornelia de Lange syndrome was detected. This result was consistent with both prenatal and postmortem examination findings, including symmetric growth restriction, a duplicated right collecting system in the fetal kidneys, and a prominent philtrum. In P27 and P28, two splice-site variants were identified in the *TSC2* gene associated with tuberous sclerosis, which often presents prenatally with cardiac rhabdomyomas as was found in these two fetuses.

Some 26 cases with positive family history of a genetic disease included in the screening panel were also analyzed. In subject P29, although no gross structural anomaly was detected on the second-trimester ultrasound exam, a pathogenic variant, c.3076C>T (p.R1026*), in the *COL1A1* gene associated with osteogenesis imperfecta type I was detected (Table 1). This variant was also present in the father, who had been previously diagnosed with osteogenesis imperfecta type I and a history of a neonatal fracture. In subjects P30–P32, paternal pathogenic variants were identified in maternal plasma cfDNA before 13 weeks gestational age. In the other five cases (P33–P37) in which fathers had a clinical diagnosis of a relevant disease, paternal pathogenic variants were not detected in the maternal plasma cfDNA. In 27 cases in which a sibling of the current fetus or a previous pregnancy was affected with a de novo pathogenic variant, the familial pathogenic variants were not detected in the current pregnancies.

Outcome information for the 76 validation (3 positive and 73 negative cases) and 422 clinical cases (32 positive and 390 negative cases) was sought via questionnaire. Providers responded to the inquiry for 233 of 498 cases (46.7%, Fig. 3). Pregnancy outcome data were obtained for 26 of 35 (74.2%) positive cases with 1 of 35 (2.9%) spontaneous abortion, 8 of 35 (22.9%) elective terminations, 7 of 35 (20%) neonatal demise, and 10 of 35 (28.6%) delivery with neonatal survival (Fig. 3). Results from invasive or postnatal sequencing tests were available for 18 of 35 (51.4%) positive cases among which 10 of 35 (28.6%) elected invasive testing and 8 of 35 (22.9%) underwent postnatal testing or testing on product of conception. In addition, 2 of 35 (5.7%) cases were confirmed by postnatal clinical examination without further testing (Fig. 3). Altogether, positive results of all 20 of 20 (100%) cases in which we could obtain diagnostic information were confirmed. Of the other positive cases, 9 of 35 (25.7%) declined any diagnostic testing. Diagnostic information for 6 of 35 (17.1%) positive cases was not available.

Table 1 | Results for positive cases with abnormal ultrasound findings and those with positive paternal history

Subject	Gestational age (weeks)	Maternal age (years)	Paternal age (years)	Prenatal finding	Fetal fraction (%)	Screening results	Variant fraction (%)	Disease	Confirmation study and pregnancy outcome[a]
P1	27	33	36	Short and bowed femur	8.2	*COL1A1*: c.1678G>A (p.G560S)	2.9	Osteogenesis imperfecta	Confirmed by the secondary assay and invasive testing
P2	20	29	31	Skeletal dysplasia	7.2	*COL1A1*: c.3425G>T (p.G1142V)	3.2	Osteogenesis imperfecta	Confirmed by the secondary assay and invasive testing, spontaneous abortion
P3	17	28	30	Bent and broken bones	20.7	*COL1A1*: c.4339delG (p.V1447Lfs*79)	8.1	Osteogenesis imperfecta	Confirmed by the secondary assay
P4	35	34	38	Skeletal dysplasia and micromelia	20.4	*COL1A1*: c.887G>C (p.G296A)	6.9	Osteogenesis imperfecta	Confirmed by the secondary assay and postnatal clinical examination, delivered
P5	30	22	34	Small thorax, bowed and short femur	13.0	*COL1A2*: c.2503G>A (p.G835S)	6.1	Osteogenesis imperfecta	Confirmed by the secondary assay
P6	26	21	34	Skeletal dysplasia	17.6	*COL1A2*: c.2684G>A (p.G895D)	7.5	Osteogenesis imperfecta	Confirmed by the secondary assay and postnatal clinical examination, neonatal demise
P7	30	41	43	Bowed and short femur	20.3	*COL1A2*: c.3034G>A (p.G1012S)	8.5	Osteogenesis imperfecta	Confirmed by the secondary assay and postnatal testing, delivered
P8	20	26	31	Skeletal dysplasia	9.3	*COL1A2*: c.2657G>T (p. G886V)	4.8	Osteogenesis imperfecta	Confirmed by the secondary assay and invasive testing
P9	22	38	45	Short femur	10.2	*FGFR3*: c.1118A>G (p.Y373C)	4.3	Thanatophoric dysplasia I	Confirmed by the secondary assay and invasive testing, neonatal demise
P10	29	33	34	Short femur and frontal bossing	22.0	*FGFR3*: c.1138G>A (p.G380R)	6.7	Achondroplasia	Confirmed by the secondary assay, delivered
P11	14	40	48	Craniomegaly, ventriculomegaly and short femur	10.8	*FGFR3*: c.1948A>G (p.K650E)	4.2	Thanatophoric dysplasia II	Confirmed by the secondary assay and invasive testing, neonatal demise
P12	19	29	29	Small thorax, short femur, abnormal skull shape and ventriculomegaly	10.5	*FGFR3*: c.1948A>G (p.K650E)	3.1	Thanatophoric dysplasia II	Confirmed by the secondary assay, neonatal demise
P13	21	27	32	Short femur	10.4	*FGFR3*: c.2419T>G (p.*807G)	3.6	Thanatophoric dysplasia I	Confirmed by the secondary assay
P14	16	35	36	Shortened long bones, hands and feet, nasal hypoplasia	9.6	*FGFR3*: c.742C>T (p.R248C)	2.7	Thanatophoric dysplasia I	Confirmed by the secondary assay, elective abortion
P15	19	44	44	Small chest-to-abdominal circumference and femur length-to-abdominal circumference ratios	11.2	*FGFR3*: c.742C>T (p.R248C)	7.7	Thanatophoric dysplasia I	Confirmed by the secondary assay, neonatal demise

Continued

Table 1 | Results for positive cases with abnormal ultrasound findings and those with positive paternal history (continued)

Subject	Gestational age (weeks)	Maternal age (years)	Paternal age (years)	Prenatal finding	Fetal fraction (%)	Screening results	Variant fraction (%)	Disease	Confirmation study and pregnancy outcome[a]
P16	22	33	38	Short femur	10.6	*FGFR3*: c.742C>T (p.R248C)	3.0	Thanatophoric dysplasia I	Confirmed by the secondary assay, neonatal demise
P17	20	36	40	Severe micromelia, short and bowed femur, concave chest and cloverleaf skull	17.6	*FGFR3*: c.742C>T (p.R248C)	5.7	Thanatophoric dysplasia I	Confirmed by the secondary assay and testing on a product of conception specimen, elective abortion
P18	30	22	27	Skeletal dysplasia	26.6	*FGFR3*: c.746C>G (p.S249C)	11.6	Thanatophoric dysplasia I	Confirmed by the secondary assay and postnatal testing, neonatal demise
P19	24	29	30	Abnormal skull shape suggestive of cranial synostosis	13.9	*FGFR2*: c.870G>T (p.W290C)	6.0	Pfeiffer syndrome	Confirmed by the secondary assay
P20	28	33	40	Craniosynostosis and abnormal hands and feet	23.2	*FGFR2:* c.758C>G (p.P253R)	8.5	Apert syndrome	Confirmed by the secondary assay and postnatal testing with clinical examination, delivered
P21	24	26	34	Ultrasound findings suggestive of Noonan spectrum disorder	15.4	*KRAS*: c.458A>T (p.D153V)	6.2	Noonan spectrum disorder	Confirmed by the secondary assay and postnatal testing, delivered
P22	13	36	40	Symmetric growth restriction, a duplicated right collecting system in the fetal kidneys and a prominent philtrum	7.4	*NIPBL*: c.1435C>T (p.R479*)	3.9	Cornelia de Lange syndrome	Confirmed by the secondary assay and testing on a product of conception specimen, elective abortion
P23	26	36	41	Intrauterine growth restriction and upper limb malformations	13.4	*NIPBL*: c.7439_7440delGA (p.R2480Lfs*5)	2.5	Cornelia de Lange syndrome	Confirmed by the secondary assay and postnatal testing and clinical examination, delivered
P24	12	33	35	Cystic hygroma and absent ductus venosus	7.2	*PTPN11*: c.188A>G (p.Y63C)	2.2	Noonan spectrum disorder	Confirmed by the secondary assay and invasive testing, delivered
P25	24	33	44	Aberrant left brachiocephalic vein and increased nuchal translucency	16.2	*RAF1*: c.782C>G (p.P261R)	6.5	Noonan spectrum disorder	Confirmed by the secondary assay and invasive testing
P26	20	36	29	Ultrasound findings suggestive of Noonan spectrum disorder	20.5	*RIT1*: c.270G>T (p.M90I)	8.3	Noonan spectrum disorder	Confirmed by the secondary assay, elective abortion
P27	31	25	27	Cardiac rhabdomyoma	19.8	*TSC2*: c.5160+2T>C	9.8	Tuberous sclerosis	Confirmed by the secondary assay, delivered
P28	28	33	31	Multiple cardiac rhabdomyomas and mild ventriculomegaly	17.0	*TSC2*: c.2220+1G>T	6.7	Tuberous sclerosis	Confirmed by the secondary assay

Continued

Table 1 | Results for positive cases with abnormal ultrasound findings and those with positive paternal history (continued)

Subject	Gestational age (weeks)	Maternal age (years)	Paternal age (years)	Prenatal finding	Fetal fraction (%)	Screening results	Variant fraction (%)	Disease	Confirmation study and pregnancy outcome[a]
P29	20	27	30	No structural anomalies, paternal history of osteogenesis imperfecta	9.0	*COL1A1*: c.3076C>T (p.R1026*)	3.2	Osteogenesis imperfecta	Confirmed by the secondary assay and postnatal testing, delivered
P30	13	32	33	Paternal diagnosis of Crouzon syndrome	17.0	*FGFR2*: c.1032G>A (p. A344A)	7.1	Crouzon syndrome	Confirmed by the secondary assay
P31	13	36	41	Paternal diagnosis of Alagille syndrome	5.4	*JAG1*: c.1446_1448delinsC (p.H483Lfs*2)	2.8	Alagille syndrome	Confirmed by the secondary assay
P32	10	26	29	Paternal diagnosis of Noonan spectrum disorder	4.9	*PTPN11*: c.836A>G (p.Y279C)	1.6	Noonan spectrum disorder	Confirmed by the secondary assay
P33	31	27	28	Normal ultrasound, father carrying *COL1A1*: c.801_802delCA (p.H267Qfs*19)	8.1	Negative	NA	NA	Confirmed by the secondary assay
P34	11	35	32	Normal ultrasound, father carrying *PTPN11*: c.1492C>T (p.R498W)	14.9	Negative	NA	NA	Confirmed by the secondary assay
P35	12	31	34	Normal ultrasound, father carrying *FGFR3*: c.1138G>A (p.G380R) and mother carrying *FGFR3*: c.1138G>C (p.G380R)	12.7	Negative[b]	N/A	N/A	Confirmed by the secondary assay and invasive testing
P36	18	28	32	Normal ultrasound, father carrying*TSC2*: c.1257+2T>C	9.3	Negative	NA	NA	Confirmed by the secondary assay
P37	20	30	33	Normal ultrasound, father carrying *COL1A1*: c.1862_1865del (p.P621Lfs*144)	19.1	Negative	NA	NA	Confirmed by the secondary assay
V1[c]	35	27	26	Micromelia skeletal dysplasia, mild dolichocephaly, small ventricular septal defect and persistent right umbilical vein	18.2	*COL1A1*: c.2164G>A (p.G722S)	9.6	Osteogenesis imperfecta	Confirmed by the secondary assay and invasive testing, delivered
V2[c]	29	31	49	Short femur and small hands	19.7	*FGFR3*: c.1138G>A (p.G380R)	7.0	Achondroplasia	Confirmed by the secondary assay and invasive testing
V3[c]	21	33	33	Pleural effusion, hydrops, cystic hygroma and small low-set ears	9.6	*RIT1*: c.229G>A (p.A77T)	3.2	Noonan spectrum disorder	Confirmed by the secondary assay and invasive testing, elective abortion

[a]Pregnancy outcome is indicated for cases in which information is available from providers. [b]The paternal variant *FGFR3*: c.1138G>A (p.G380R) was not detected whereas the assay was not informative for the fetal genotype regarding the maternal variant *FGFR3*: c.1138G>C (p.G380R). [c]Three positive cases collected during the validation study were included in the follow-up study.

Providers were also contacted and encouraged to report any false-negative results, but no such cases were reported. Providers responded to the inquiry for 198 of 463 (42.7%) negative cases (Fig. 3). For those, we determined that 127 cases yielded true-negative screening results based on diagnostic genetic testing or postnatal clinical examination. However, potential false-negative results cannot be completely ruled out owing to possible late-onset or mild phenotype. Overall, in cases in which we could obtain the clinical outcome data, we confirmed 20 true-positive, 127 true-negative, zero false-positive and zero-false negative results in the clinical follow-up study (Fig. 3 and Supplementary Table 8).

The technical advances we made in this work for non-invasive prenatal sequencing of frequent human dominant diseases include molecular indexing of cfDNA for accurate DNA molecule counting and low-level fetal variant calling; target enrichment by probe hybridization rather than PCR amplification to avoid allele dropout in highly fragmented cfDNA; full-gene sequencing for the detection of essentially all sequence mutations in the targeted genes; an integrated analysis for fetal fraction based on transmitted SNPs; and a *Z*-value-based statistical method integrating fetal fraction and sequence read counts for the filtering of low-level fetal variants (Fig. 1 and Supplementary Table 9). The presented comprehensive assay methodologies and analytical algorithms can be used to expand the current test to include many more dominant diseases.

Some monogenic diseases with relatively high incidence are often suspected when pregnancies have abnormal cardiac, skeletal, or other structural ultrasound findings. For instance, Noonan spectrum disorders are a group of autosomal dominant diseases that can have overlapping prenatal sonographic features with trisomy 21, 18, and 13 such as increased nuchal translucency in the first trimester[24,25]. *FGFR3*-related skeletal disorders are another group of diseases that are often suspected when there are shortened femurs and humeri, and/or frontal bossing on a second-trimester ultrasound exam[26,27]. For such pregnancies with abnormal prenatal ultrasound findings, an invasive diagnostic procedure followed by molecular genetic testing is often needed to clarify whether the fetus is affected. However, when a specific diagnosis is not indicated and only 'soft markers' are present (for example, increased nuchal translucency or shortened long bones), patients are often reluctant to undergo an invasive procedure and prefer non-invasive assessment.

A common molecular etiology for dominant diseases such as Noonan spectrum disorders and *FGFR3*-related skeletal disorders includes de novo variants, which occur more frequently with advanced paternal age[28]. In 28 positive clinical cases in which we identified de novo pathogenic or likely pathogenic variants, the average paternal age was ~36 years. Current guidelines for counseling for advanced paternal age use age 40 as an arbitrary cutoff[29,30], which was four years older than the mean paternal age of those having an affected fetus with a de novo variant in this study. This suggests that further systematic studies may be warranted to establish a more precise advanced paternal age cutoff for genetic counseling and potential high-risk screening. Alternatively, such studies may clarify whether screening should be more widely offered, irrespective of paternal age.

Overall, the NIPS approach presented here provides valuable molecular information regarding the fetal risks associated with frequent human dominant monogenic disorders and helps guide physician and parent decisions for further evaluation and management of the pregnancy.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-018-0334-x.

Received: 15 May 2018; Accepted: 17 December 2018;
Published online: 28 January 2019

References

1. Norton, M. E. et al. Cell-free DNA analysis for non-invasive examination of trisomy. *N. Engl. J. Med.* **372**, 1589–1597 (2015).
2. Dar, P. et al. Clinical experience and follow-up with large scale single-nucleotide polymorphism-based non-invasive prenatal aneuploidy testing. *Am. J. Obstet. Gynecol.* **211**, 527.e1–527.e17 (2014).
3. Baird, P. A., Anderson, T. W., Newcombe, H. B. & Lowry, R. B. Genetic disorders in children and young adults: a population study. *Am. J. Hum. Genet.* **42**, 677–693 (1988).
4. Lo, Y. M. et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for non-invasive prenatal diagnosis. *Am. J. Hum. Genet.* **62**, 768–775 (1998).
5. Chiu, R. W. et al. Non-invasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc. Natl Acad. Sci. USA* **105**, 20458–20463 (2008).
6. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Non-invasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc. Natl Acad. Sci. USA* **105**, 16266–16271 (2008).
7. Lo, Y. M. et al. Presence of fetal DNA in maternal plasma and serum. *Lancet* **350**, 485–487 (1997).
8. McCullough, R. M. et al. Non-invasive prenatal chromosomal aneuploidy testing—clinical experience: 100,000 clinical samples. *PLoS One* **9**, e109173 (2014).
9. Petersen, A. K. et al. Positive predictive value estimates for cell-free non-invasive prenatal screening from data of a large referral genetic diagnostic laboratory. *Am. J. Obstet. Gynecol.* **217**, 691.e1–691.e6 (2017).
10. Chitty, L. S. & Lo, Y. M. Non-invasive prenatal screening for genetic diseases using massively parallel sequencing of maternal plasma DNA. *Cold Spring Harb. Perspect. Med.* **5**, a023085 (2015).
11. Agatisa, P. K. et al. A first look at women's perspectives on non-invasive prenatal testing to detect sex chromosome aneuploidies and microdeletion syndromes. *Prenat. Diagn.* **35**, 692–698 (2015).
12. Lo, K. K. et al. Limited clinical utility of non-invasive prenatal testing for subchromosomal abnormalities. *Am. J. Hum. Genet.* **98**, 34–44 (2016).
13. Rowley, P. T., Loader, S. & Kaplan, R. M. Prenatal screening for cystic fibrosis carriers: an economic evaluation. *Am. J. Hum. Genet.* **63**, 1160–1174 (1998).
14. Nelson, W. B., Swint, J. M. & Caskey, C. T. An economic evaluation of a genetic screening program for Tay-Sachs disease. *Am. J. Hum. Genet.* **30**, 160–166 (1978).
15. Yang, Y. et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* **312**, 1870–1879 (2014).
16. Chitty, L. S. 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.* **35**, 656–662 (2015).
17. You, Y. et al. Integration of targeted sequencing and NIPT into clinical practice in a Chinese family with maple syrup urine disease. *Genet. Med.* **16**, 594–600 (2014).
18. Lench, N. et al. The clinical implementation of non-invasive prenatal diagnosis for single-gene disorders: challenges and progress made. *Prenat. Diagn.* **33**, 555–562 (2013).
19. Meacham, F. et al. Identification and correction of systematic error in high-throughput sequence data. *BMC Bioinformatics* **12**, 451 (2011).
20. Loman, N. J. et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* **30**, 434–439 (2012).
21. Schirmer, M. et al. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* **43**, e37 (2015).
22. Newman, A. M. et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* **34**, 547–555 (2016).
23. Samango-Sprouse, C. et al. SNP-based non-invasive prenatal testing detects sex chromosome aneuploidies with high accuracy. *Prenat. Diagn.* **33**, 643–649 (2013).
24. Roberts, A. E., Allanson, J. E., Tartaglia, M. & Gelb, B. D. Noonan syndrome. *Lancet* **381**, 333–342 (2013).
25. Nisbet, D. L., Griffin, D. R. & Chitty, L. S. Prenatal features of Noonan syndrome. *Prenat. Diagn.* **19**, 642–647 (1999).
26. Vajo, Z., Francomano, C. A. & Wilkin, D. J. The molecular and genetic basis of fibroblast growth factor receptor 3 disorders: the achondroplasia family of skeletal dysplasias, Muenke craniosynostosis, and Crouzon syndrome with acanthosis nigricans. *Endocr. Rev.* **21**, 23–39 (2000).
27. Krakow, D., Lachman, R. S. & Rimoin, D. L. Guidelines for the prenatal diagnosis of fetal skeletal dysplasias. *Genet. Med.* **11**, 127–133 (2009).
28. Goriely, A. & Wilkie, A. O. Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. *Am. J. Hum. Genet.* **90**, 175–200 (2012).

29. Friedman, J. M. Genetic disease in the offspring of older fathers. *Obstet. Gynecol.* **57**, 745–749 (1981).
30. Toriello, H. V., Meck, J. M., Professional, P. & Guidelines, C. Statement on guidance for genetic counseling in advanced paternal age. *Genet. Med.* **10**, 457–460 (2008).

Acknowledgements

Baylor Genetics Laboratories and Baylor College of Medicine provided funds to support this study. A.K.M. was supported by institutional funds from Baylor College of Medicine. K.W.C. was partially supported by Vice-Chancellor Discretionary Fund for CUHK-Baylor College of Medicine Joint Centre for Medical Genetics. We thank all contributing healthcare providers for their work and support on this study.

Author Contributions

A.B., J.Z., C.M.E., and L.-J.W. designed the study; J.L., Y.F., J.S., S.C., H.D., X.G., and G.W. conducted the experiments; J.Z., J.L., Y.F., Y.J., E.S.S., S.P., S.C., H.D., X.G., G.W., C.A.S., H.M., A.B., F.X., Y.Y., A. Purgason, A. Pourpak, Z.C., X.W., Y.W., S.K., K.W.C., R.J.W., J.B.S., S.P., A.K.M., I.B.V.d.V., A.B., L.-J.W., and C.M.E. conducted the validation and/or clinical data analyses; C.A.S., J.L., Y.F., and J.Z. conducted the statistical analyses; J.Z. and J.L. wrote the manuscript; J.Z., L.-J.W., and C.M.E. supervised the project.

Competing interests

The joint venture of Department of Molecular and Human Genetics at Baylor College of Medicine (BCM) and Baylor Genetics Laboratories (BG) derives revenue from the clinical sequencing offered at BG and the authors who are BCM faculty members or BG employees are indicated in the affiliation section. J.B.S. and S.P. are employees of Natera who provided samples for validating fetal fraction calculation and contributed to the clinical data collection and analysis for the outcome study. The patent application related to this work has been filed (WO2018049049A1) by BCM and BG Laboratories, including laboratory methods of non-invasive prenatal testing to detect dominant monogenic disorders.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-018-0334-x.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-018-0334-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.Z.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Ethics statement. This study has complied with the protocols approved by the Institutional Review Board for Human Subject Research at the Baylor College of Medicine. Informed consent was obtained for genetic testing performed in a Clinical Laboratory Improvement Amendments (CLIA) and College of American Pathologist (CAP) accredited laboratory. The clinical test was approved by the New York State Department of Health according to the requirements for laboratory-developed tests.

Study participants and clinical testing. From October 2016 through April 2018, a total of 76 validation samples and 422 clinical samples that met the acceptance criteria were tested and the clinical results were retrospectively analyzed. For the clinical test, we only accepted samples collected from singleton pregnancies with a gestational age above nine weeks. We required two tubes of maternal venous blood (a minimum of 8 ml in each tube) collected in BCT Streck tubes and arriving at our laboratory within 5 d of sample collection. To assure the test performance, we reported only test results for samples with fetal fraction ≥4.5% and with minimum sequencing coverage of 200× for ≥97% of target regions. Paternal blood or saliva samples were also collected for the analysis to facilitate variant interpretation. Clinical information such as pregnancy history, fetal ultrasound findings, and medical and family history of parents were collected.

Cell-free DNA extraction. The maternal plasma was separated through a two-step centrifugation process[31]. In the first step, plasma was separated from the whole blood by centrifugation of the collection tube at 1,600*g* for 15 min at 4 °C. The second step was centrifugation of the recovered plasma at 16,000*g* for 10 min at 4 °C. The cfDNA was then extracted using a QIAamp circulating nucleic acid extraction kit (Qiagen) or MagMAX cfDNA isolation kit from Thermo Fisher. The cfDNA was eluted in 50 µl of elution buffer. Total genomic DNA was extracted from maternal buffy coat or paternal leukocytes according to the manufacturer's protocol (Chemagen).

Library preparation, unique molecular indexing and next-generation sequencing. After cfDNA extraction, individual DNA molecules were tagged using Y-shaped adapters conjugated to a UMI comprising six degenerate nucleotides. Each sample was then indexed with a unique sample barcode using PCR. A previously described protocol for hybridization-based enrichment was used for NGS library preparation[32]. Sequencing was performed on a HiSeq 2500 instrument (Illumina) using 2 × 100 paired-end sequencing in the rapid mode. There are potentially 4^{14} or 2.7×10^8 combinations of different indexed adaptors that can be incorporated in each end of library inserts. Because the UMIs were used at a ~10^5 molar excess over the amount of the cfDNA fragments derived from ~10 ng input cfDNA, each insert can be individually labeled with a different UMI (Fig. 1a).

Next-generation sequencing data analysis. After demultiplexing, NGS reads that share the same UMI were grouped together. Consensus reads were consolidated for each subgroup and the base call at a particular locus was made when the majority (≥66.6%) of the consensus reads contained the same allele. A 'no call' was given if 33.3–66.6% of the reads contained a variant nucleotide (Fig. 1b). The consolidated reads were used for alignment to the human genome hg19 reference sequence followed by variant calling. Variants with a minor allele frequency ≥0.5% were called by using the NextGENe software version 2.3 (SoftGenetics). Classification of genetic variants was conducted by trio analysis using the sequencing data collected from maternal plasma cfDNA and parental genomic DNA according to guidelines issued by American College of Medical Genetics and Genomics[33].

Fetal fraction calculation and standard normal transformation for fetal variants. The FF was calculated from the average of two populations of informative loci in SNPs (Supplementary Table 2). When the mother is homozygous for a reference allele and the fetus inherits an alternative allele from the father at the same locus, the expected variant allele fraction in the total cfDNA (VAFp) is half the FF. Similarly, when the mother is homozygous for an alternative allele and the fetus inherits the paternal reference allele, the expected variant allele fraction (VAFm) is 1 – (0.5*FF). The FF is calculated based on the average of all informative transmitted SNP loci using the following formula: $FF = VAFp_{average} + (1 - VAFm_{average})$. For each called variant, a *Z*-value is calculated for the purpose of standard normal transformation using the following formula: $Z = (UMI_{observed} - UMI_{expected})/UMI_{expected}$ where *UMI* is the unique molecular index used to label individual DNA molecules; $UMI_{observed}$ is the observed number of different *UMI*s conjugated with distinct DNA fragments harboring the variant allele at a particular locus and $UMI_{expected}$ is the expected number of UMIs associated with a fetal heterozygous variant as a function of the fetal fraction and the total sequencing depth at the variant locus (represented by UMI_{total}): $UMI_{expected} = UMI_{total} * FF/2$.

Confirmation of positive results identified by primary capture–based NGS assay. All pathogenic and likely pathogenic variants found by the primary capture–based NGS assay were confirmed by a secondary amplicon–based NGS assay using cfDNA as template from a separate tube of maternal blood. First gene-specific primers were used to amplify DNA fragments containing the variant of interest with an amplicon size ≤125 bp. In the second amplification Illumina adaptors for sequencing were introduced by PCR. The condition of the first PCR was: 95 °C, 5 min; 10 cycles of 95 °C for 15 s, 65 °C for 15 s, 72 °C for 30 s with –1.0° per cycle; followed by 30 cycles of 95 °C for 15 s, 55 °C for 15 s, 72 °C for 30 s; 72 °C for 5 min and holding at 4 °C. The condition of the second PCR was: 95 °C for 45 s; 10 cycles of 98 °C for 15 s, 60 °C for 30 s, 72 °C for 30 s; 72 °C for 1 min; and holding at 4 °C. The secondary confirmation assay was performed concurrently for both the parental leukocyte DNA and maternal plasma cfDNA to examine whether there was low-level maternal mosaicism or paternal inheritance. When an invasive specimen (for example amniotic fluid) or a postnatal sample was available, Sanger sequencing was used to confirm the positive cfDNA finding.

Statistical analysis. Statistical analyses were performed using the unpaired Student's *t*-test for comparison of UMI- and non-UMI-based NGS variant calling errors ($n = 4$, degrees of freedom = 6) in loci with different variant-allele fractions. Data were considered significant when the two-tailed *P* values were <0.05 (95% confidence interval) and were considered nonsignificant when *P* values were ≥0.05. Pearson correlation coefficient was calculated to show the correlation of fetal fraction results calculated by the in-house developed SNP-based methods and other methods.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability

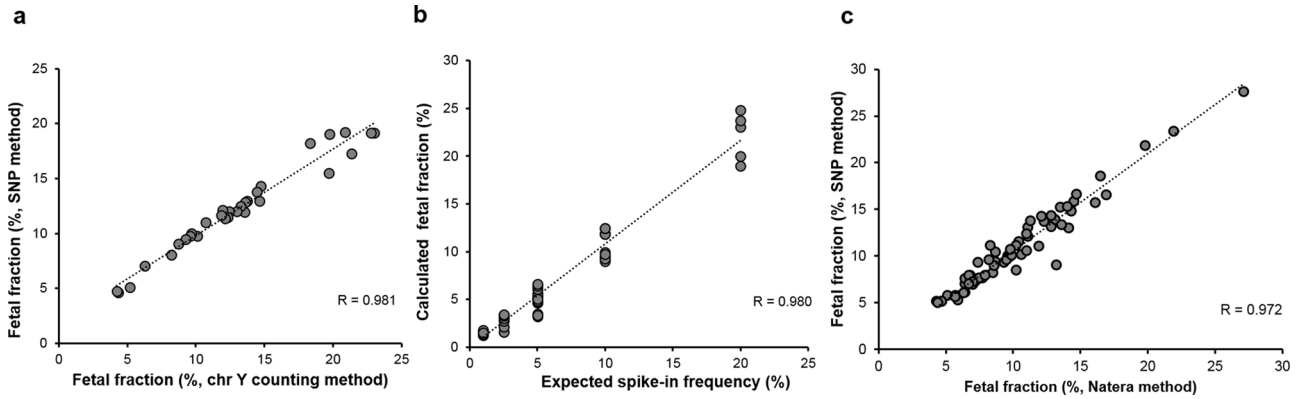
The customized script for the UMI-based deduplication of the NGS reads can be found at https://sourceforge.net/projects/BGNIPS.

Data availability

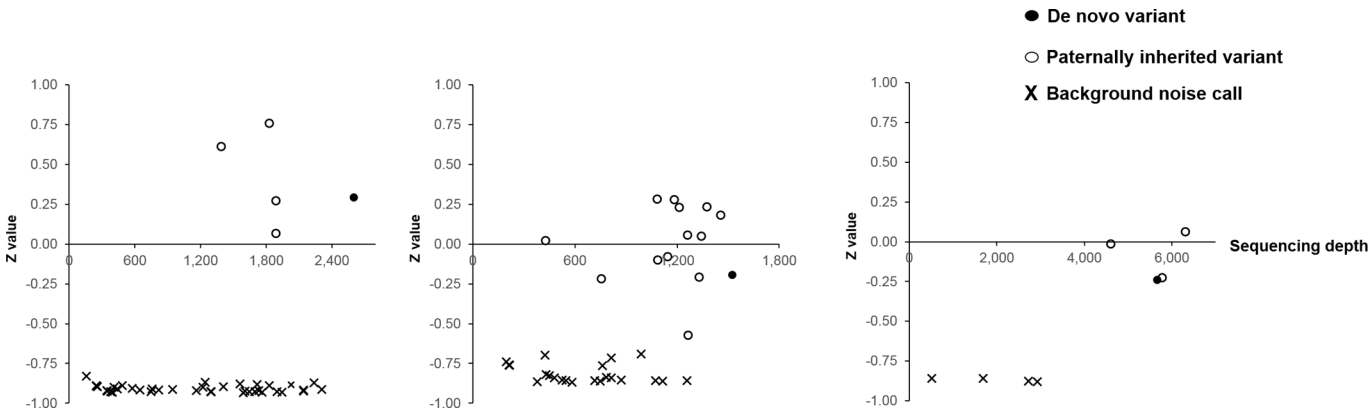
These authors declare that all essential data supporting the conclusion of the study as well as detailed assay protocols, analytical algorithms, and customized computational codes are within the paper and supplementary materials. All the disease-causing variants and the key phenotypes found in the subjects can be found at the ClinVar database (http://www.ncbi.nlm.nih.gov/clinvar/variation/) with accession numbers SCV000854595–SCV000854628. Subjects' identifiable information (including their genomic sequencing data) is kept in our clinical laboratory, which is a CLIA and CAP certified laboratory and a HIPAA-compliant environment, to protect subjects' privacy. Non-identifiable sequencing data (for example, individual variant sequencing data generated by locus-specific sequencing) can be provided on request from the authors. Source data for Fig. 1 and Extended Data Figs. 1 and 2 are available online.

References

31. Breitbach, S. et al. Direct quantification of cell-free, circulating DNA from unpurified plasma. *PLoS One* **9**, e87838 (2014).
32. Yang, Y. et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* **369**, 1502–1511 (2013).
33. 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.* **17**, 405–424 (2015).



Extended Data Fig. 1 | Validation of a SNP-based fetal fraction calculation method. a, The comparison of fetal fraction calculation between our SNP-based method and a Y chromosome marker method. A total of 31 samples from pregnancies with male fetuses were used. **b**, The comparison of the calculated and expected fetal fraction in the spike-in experiment. Five sets of spike-in samples were created by mixing a proband's DNA with maternal DNA to mimic fetal fraction ranging from 1 to 20%. A total of 43 spike-in samples were used. **c**, Comparison of the fetal fraction calculated for 67 samples collected from pregnant women (fetal fraction ranging 5.0–27.6%) by our method and an independent method developed by another laboratory. Pearson correlation coefficient is shown as the *R* value calculated by Microsoft Excel.



Extended Data Fig. 2 | *Z*-value plots for three representative samples which had de novo, paternally inherited, and false analytical variants with UMI deduplication and consolidation for NGS reads and variant calling/filtering. The paternally inherited variants and confirmed de novo variants deemed true positives had $Z > -0.6$.

nature research

Corresponding author(s): Jinglan Zhang

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect the data.
Data analysis	The customized script for the UMI-based deduplication of the NGS reads is available upon request or can be found at https://sourceforge.net/projects/BGNIPS. NextGENe software version 2.3 (SoftGenetics, State College, PA) was used for sequence alignment and variant calling.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

These authors declare that all essential data supporting the conclusion of the study as well as detailed assay protocols, analytical algorithms and customized

computational codes are within the paper and supplementary materials. All the disease-causing variants and the key phenotypes found in the patients can be found at the ClinVar database (http://www.ncbi.nlm.nih.gov/clinvar/variation/) with the accession numbers SCV000854595 - SCV000854628. Patients' identifiable information (including their genomic sequencing data) is kept in our clinical laboratory which is a CLIA and CAP certified laboratory and a HIPAA-compliant environment, in order to protect patients' privacy. Non-identifiable sequencing data (e.g., individual variant sequencing data generated by locus-specific sequencing) can be provided upon request from the authors. Source data for Figure 1, Extended Data Figure 1 and 2 are available online.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	There were 76 samples (with 554 true positive variants and 8,038,7922 true negative loci) included in the validation study and 422 samples (147 with outcome data) included in the clinical study. These samples were included to calculate the test sensitivity and specificity. The sample size we chose allowed a probability of 95% or above to observe a possible measuring error at the variant level for the analytical validation and a probability of 90% or above to observe a possible measuring error at the case level for the clinical study.
Data exclusions	There were no data exclusion for our analytical validation study. For our clinical samples, we rejected samples that did not meet our sample intake requirements as they were not collected in the same condition as our validation samples. Of these samples, 18 were excluded because they did not meet the minimal sequencing data coverage or fetal fraction cut-off at 4.5%. Eleven samples were rejected due to handling issues such as insufficient blood volume, delayed shipment, missing paternal samples and incomplete sample identification information. Three samples were excluded because maternal pathogenic variants were identified. Two were rejected because they originated from women with twin pregnancies and another two from women with pregnancies conceived with donor eggs for whom DNA from the biological mother was not available. The clinical sample acceptance or exclusion criteria were predetermined based on the standard operation procedure established during the method validation. For the clinical outcome study, we collected data for patients which the providers responded to our inquiry and we only used those patients with available diagnostic data for the calculation of clinical sensitivity and specificity.
Replication	For all positive findings identified in the primary assay by the capture-based NGS method, a confirmatory test was performed by a second method using targeted amplicon sequencing with the cell-free DNA extracted from a second blood draw of the mother. In addition, another confirmation test was performed by using an invasive fetal specimen or a postnatal specimen whenever such samples were available to us. All positive results in our clinical samples were confirmation by the replication studies.
Randomization	This is not relevant to our study. We accepted all samples as long as they were collected in the same conditions predetermined during our validation studies.
Blinding	For all of our clinical samples, we report positive findings based on the sequencing studies using the maternal cell-free DNA. The clinical confirmation study using an invasive fetal specimen or a postnatal specimen was performed at a later time so we would not know the outcome of fetus when we performed the primary test on the maternal cell-free DNA.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	No unique materials were used. All synthetic materials (PCR primers, probes and NGS library construction adaptors) were

Obtaining unique materials	manufactured by commercial vendors (Integrated DNA Technologies, Roche, and SigmaAldrich) according to their standard protocols. Information regarding how these materials were designed are within in the paper.

Human research participants

Policy information about studies involving human research participants

Population characteristics	This study included patients seeking prenatal diagnosis or genetic disease risk assessment for their pregnancies due to family history, prenatal ultrasound findings, previous screening results, advanced paternal or maternal age or parental concerns. These patients represent a diverse population from different ethnicity or childbearing age groups. We have not observed any covariate-relevant population characteristics.
Recruitment	This study was approved by the Institutional Review Board for Human Subject Research at the Baylor College of Medicine (Houston, TX). Informed consent was obtained for clinical genetic testing performed in a CLIA (Clinical Laboratory Improvement Amendments) and CAP (College of American Pathologist) accredited laboratory. From October 2016 through April 2018, a total of 76 validation samples and 422 clinical samples were tested for the evaluation of fetal risks associated with monogenic disorders and the clinical results were retrospectively analyzed. This cohort consists of consecutive samples collected from a diverse population with different ethnicity, childbearing age and a wide range of testing indications. We have not observed any self-selection bias or other biases that may impact the results of this study.