# Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood

H. Christina Fan*, Yair J. Blumenfeld†, Usha Chitkara†, Louanne Hudgins‡, and Stephen R. Quake*§

*Department of Bioengineering, Stanford University and Howard Hughes Medical Institute, 318 Campus Drive, Clark Center, Room E300, Stanford, CA 94305; †Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Stanford University, 300 Pasteur Drive, Room HH333, Stanford, CA 94305; and ‡Division of Medical Genetics, Department of Pediatrics, Stanford University, 300 Pasteur Drive, Stanford, CA 94305

Communicated by Leonard A. Herzenberg, Stanford University School of Medicine, Stanford, CA, August 22, 2008 (received for review July 13, 2008)

**We directly sequenced cell-free DNA with high-throughput shotgun sequencing technology from plasma of pregnant women, obtaining, on average, 5 million sequence tags per patient sample. This enabled us to measure the over- and underrepresentation of chromosomes from an aneuploid fetus. The sequencing approach is polymorphism-independent and therefore universally applicable for the noninvasive detection of fetal aneuploidy. Using this method, we successfully identified all nine cases of trisomy 21 (Down syndrome), two cases of trisomy 18 (Edward syndrome), and one case of trisomy 13 (Patau syndrome) in a cohort of 18 normal and aneuploid pregnancies; trisomy was detected at gestational ages as early as the 14th week. Direct sequencing also allowed us to study the characteristics of cell-free plasma DNA, and we found evidence that this DNA is enriched for sequences from nucleosomes.**

fetal DNA | next-generation sequencing | noninvasive prenatal diagnosis | Down syndrome | trisomy

Fetal aneuploidy and other chromosomal aberrations affect 9 of 1,000 live births (1). The gold standard for diagnosing chromosomal abnormalities is karyotyping of fetal cells obtained via invasive procedures such as chorionic villus sampling and amniocentesis. These procedures impose small but potentially significant risks to both the fetus and the mother (2). Noninvasive screening of fetal aneuploidy using maternal serum markers and ultrasound are available but have limited reliability (3–5). There is therefore a desire to develop noninvasive genetic tests for fetal chromosomal abnormalities.

Since the discovery of intact fetal cells in maternal blood, there has been intense interest in trying to use them as a diagnostic window into fetal genetics (6–9). Although this has not yet moved into practical application (10), the later discovery that significant amounts of cell-free fetal nucleic acids also exist in maternal circulation has led to the development of new noninvasive prenatal genetic tests for a variety of traits (11, 12). However, measuring aneuploidy remains challenging because of the high background of maternal DNA; fetal DNA often constitutes $<10\%$ of total DNA in maternal cell-free plasma (13). Recently developed methods for aneuploidy detection focus on allelic variation between the mother and the fetus. Lo *et al.* (14) demonstrated that allelic ratios of placental-specific mRNA in maternal plasma could be used to detect trisomy 21 (T21) in certain populations. Similarly, they also showed the use of allelic ratios of imprinted genes in maternal plasma DNA to diagnose trisomy 18 (T18) (15). Dhallan *et al.* (16) used fetal-specific alleles in maternal plasma DNA to detect trisomy 21. However, these methods are limited to specific populations because they depend on the presence of genetic polymorphisms at specific loci. We and others argued that it should be possible, in principle, to use digital PCR to create a universal, polymorphism-independent test for fetal aneuploidy by using maternal plasma DNA (17–19), but because of technical challenges relating to the low fraction of fetal DNA, such a test has not yet been practically realized.

An alternative method to achieve digital quantification of DNA is direct shotgun sequencing, followed by mapping to the chromosome of origin and enumeration of fragments per chromosome. Recent advances in DNA-sequencing technology allow massively parallel sequencing (20), producing tens of millions of short sequence tags in a single run and enabling a deeper sampling than can be achieved by digital PCR. By counting the number of sequence tags mapped to each chromosome, the over- or underrepresentation of any chromosome in maternal plasma DNA contributed by an aneuploid fetus can be detected. This method does not require the differentiation of fetal versus maternal DNA, and with large enough tag counts, it can be applied to arbitrarily small fractions of fetal DNA. We demonstrate here the successful use of shotgun sequencing to detect fetal trisomy 21 (Down syndrome), trisomy 18 (Edward syndrome), and trisomy 13 (T13) (Patau syndrome) noninvasively by using cell-free fetal DNA in maternal plasma. This forms the basis of a universal, polymorphism-independent noninvasive diagnostic test for fetal aneuploidy. The sequence data also allowed us to characterize plasma DNA in unprecedented detail, suggesting that it is enriched for nucleosome-bound fragments.

## Results

**Shotgun Sequencing of Cell-Free Plasma DNA.** Cell-free plasma DNA from 18 pregnant women and a male donor, as well as whole-blood genomic DNA from the same male donor, were sequenced on the Solexa/Illumina platform. We obtained on average ≈10 million 25-bp sequence tags per sample. Approximately 50% (i.e., ≈5 million) of the reads mapped uniquely to the human genome with, at most, one mismatch against the human genome, covering ≈4% of the entire genome. An average of ≈154,000, ≈135,000, and ≈65,700 sequence tags mapped to chromosomes 13, 18, and 21, respectively. The number of sequence tags for each sample is detailed in supporting information (SI) Table S1.

We observed a nonuniform distribution of sequence tags across each chromosome. This pattern of intrachromosomal variation was common among all samples, including randomly sheared genomic DNA, indicating that the observed variation was most probably due to sequencing artifacts. We applied a sliding window of 50 kb across each chromosome and counted the number of tags falling within each window. The median count per 50-kb window for each chromosome was selected. The median of the autosomal values was used as a normalization constant to account for the differences in

Author contributions: H.C.F., Y.J.B., U.C., L.H., and S.R.Q. designed research; H.C.F. performed research; H.C.F. analyzed data; Y.J.B., U.C., and L.H. designed the IRB-approved clinical protocol and coordinated patient recruitment and enrollment; and H.C.F., Y.J.B., and S.R.Q. wrote the paper.

Conflict of interest statement: S.R.Q. is a founder, shareholder, and consultant of Fluidigm Corporation. S.R.Q. and H.C.F. have applied for a patent relating to the method described in this study. Other authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: Sequence data have been deposited at the National Center for Biotechnology Information short read archive (www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi), accession no. SRA001174.

§To whom correspondence should be addressed. E-mail: quake@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808319105/DCSupplemental.

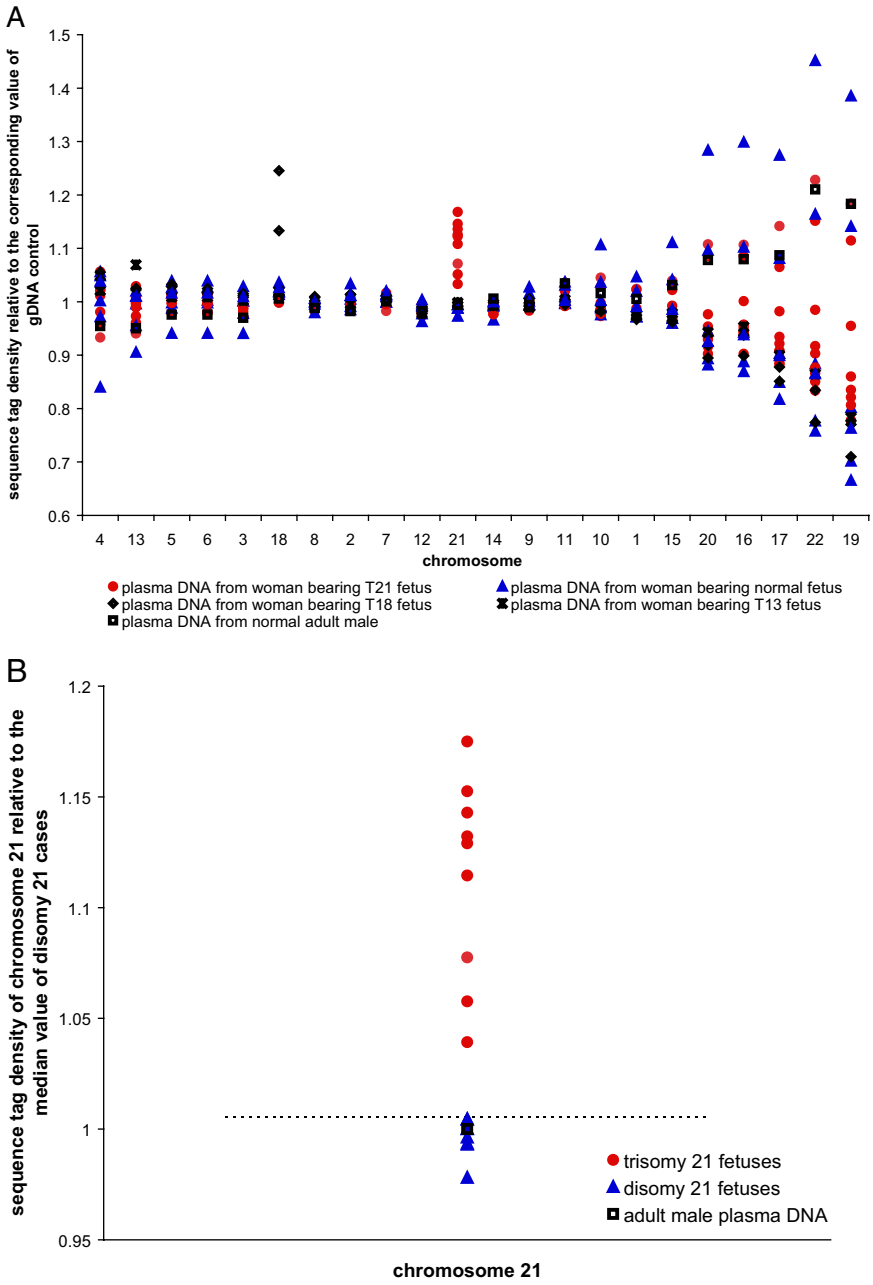


**Fig. 1.** Fetal aneuploidy is detectable by the overrepresentation of the affected chromosome in maternal blood. (*A*) Sequence tag density relative to the corresponding value of genomic DNA control; chromosomes are ordered by increasing GC content. (*B*) Chromosome 21 sequence tag density relative to the median chromosome 21 sequence tag density of the normal cases. Note that the values of three disomy 21 cases overlap at 1.0. The dashed line represents the upper boundary of the 99% confidence interval constructed from all disomy 21 samples. Number of disomy 21 samples = 9. Number of trisomy 21 samples = 9.

total number of sequence tags obtained for different samples. (From this point forward, "sequence tag density" refers to the normalized value and is used for comparing different samples and for subsequent analysis). The interchromosomal variation within each sample was also consistent among all samples (including genomic DNA control). The mean sequence tag density of each chromosome correlates with the GC content of the chromosome ($P < 10^{-9}$) (Fig. S1 *A* and *B*). The standard deviation of sequence tag density for each chromosome also correlates with the absolute degree of deviation in chromosomal GC content from the genome-wide GC content ($P < 10^{-12}$) (Fig. S1 *A* and *C*). The GC content of sequenced tags of all samples (including the genomic DNA control) was, on average, ≈10% higher than the value of the sequenced human genome (41%) (21) (Table S1), suggesting that there is a strong GC bias stemming from the sequencing process. We plotted in Fig. 1*A* the sequence tag density for each chromosome (ordered by increasing GC content) relative to the corresponding value of the genomic DNA control to remove such bias.

**Detection of Fetal Aneuploidy.** The distribution of chromosome 21 sequence tag density for all nine T21 pregnancies is clearly separated from that of pregnancies bearing disomy 21 fetuses ($P < 10^{-5}$, Student's *t* test) (Fig. 1 *A* and *B*). The coverage of chromosome 21 for T21 cases is ≈4–18% higher (average ≈11%) than that of the disomy 21 cases. Because the sequence tag density of chromosome 21 for T21 cases should be $(1 + \varepsilon/2)$ of that of disomy 21 pregnancies, where $\varepsilon$ is the fraction of total plasma DNA originating from the fetus (see *SI Appendix* for derivations), such increase in chromosome 21 coverage in T21 cases corresponds to a fetal DNA fraction of ≈8–35% (average ≈23%) (Table S1 and Fig. 2). We constructed a 99% confidence interval of the distribution of chromosome 21 sequence tag density of disomy 21 pregnancies. The values for all nine T21 cases lie outside the upper boundary of the confidence interval, and those for all nine disomy 21 cases lie below the boundary (Fig. 1*B*). If we used the upper bound of the confidence interval as a threshold value for detecting T21, the minimum fraction of fetal DNA that would be detected is ≈2%.

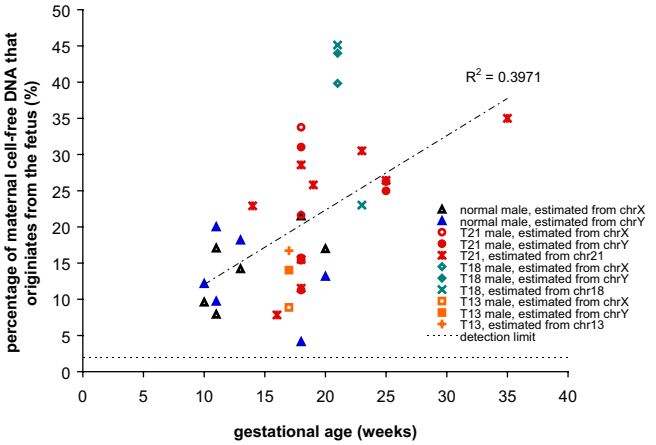


**Fig. 2.** Fetal DNA fraction and gestational age. The fraction of fetal DNA in maternal plasma correlates with gestational age. Fetal DNA fraction was estimated in three different ways: (*i*) from the additional amount of chromosomes 13, 18, and 21 sequences for T13, T18, and T21 cases, respectively; (*ii*) from the depletion in amount of chromosome X sequences for male cases; (*iii*) from the amount of chromosome Y sequences present for male cases. The horizontal dashed line represents the estimated minimum fetal DNA fraction required for the detection of aneuploidy. For each sample, the values of fetal DNA fraction calculated from the data of different chromosomes were averaged. There is a statistically significant correlation between the average fetal DNA fraction and gestational age ($P = 0.0051$). The dashed line represents the simple linear regression line between the average fetal DNA fraction and gestational age. The $R^2$ value represents the square of the correlation coefficient.

Plasma DNA of pregnant women carrying T18 fetuses (two cases) and a T13 fetus (one case) were also directly sequenced. Overrepresentation was observed for chromosomes 18 and 13 in T18 and T13 cases, respectively (Fig. 1*A*). Although there were not enough positive samples to measure a representative distribution, it is encouraging that all of these three positives are outliers from the distribution of disomy values. The T18 are large outliers and are clearly statistically significant ($P < 10^{-7}$), whereas the statistical significance of the single T13 case is marginal ($P < 0.05$). Fetal DNA fraction was also calculated from the overrepresented chromosome as described above (Fig. 2 and Table S1).

**Fetal DNA Fraction in Maternal Plasma.** Using digital TaqMan PCR for a single locus on chromosome 1, we estimated the average cell-free DNA concentration in the sequenced maternal plasma samples to be ≈360 cell equivalents per milliliter of plasma (range: 57–761 cell equivalents per milliliter of plasma) (Table S1), in rough accordance with previously reported values (13). The cohort included 12 male pregnancies (6 normal cases, 4 T21 cases, 1 T18 case, and 1 T13 case) and 6 female pregnancies (5 T21 cases and 1 T18 case). DYS14, a multicopy locus on chromosome Y, was detectable in maternal plasma by real-time PCR in all these pregnancies but not in any of the female pregnancies (data not shown). The fraction of fetal DNA in maternal cell-free plasma DNA is usually determined by comparing the amount of fetal-specific locus (such as the SRY locus on chromosome Y in male pregnancies) to that of a locus on any autosome that is common to both the mother and the fetus by using quantitative real-time PCR (13, 22, 23). We applied a similar duplex assay on a digital PCR platform (see *Materials and Methods*) to compare the counts of the SRY locus and a locus on chromosome 1 in male pregnancies. SRY locus was not detectable in any plasma DNA samples from female pregnancies. We found with digital PCR that for the majority samples, fetal DNA constituted ≤10% of total DNA in maternal plasma (Table S1), agreeing with previously reported values (13).

The percentage of fetal DNA among total cell-free DNA in maternal plasma can also be calculated from the density of sequence tags of the sex chromosomes for male pregnancies. By comparing the sequence tag density of chromosome Y of plasma DNA from male pregnancies to that of adult male plasma DNA, we estimated fetal DNA percentage to be, on average, ≈19% (range: 4–44%) for all male pregnancies (Table S1 and Fig. 2). Because human males have one fewer chromosome X than human females, the sequence tag density of chromosome X in male pregnancies should be $(1 - \varepsilon/2)$ of that of female pregnancies, where $\varepsilon$ is fetal DNA fraction (see *SI Appendix* for derivation). We indeed observed underrepresentation of chromosome X in male pregnancies as compared with that of female pregnancies (Fig. S2). Based on the data from chromosome X, we estimated fetal DNA percentage to be, on average, ≈19% (range: 8–40%) for all male pregnancies (Table S1 and Fig. 2). The fetal DNA percentage estimated from chromosomes X and Y for each male pregnancy sample correlated with each other ($P = 0.0015$) (Fig. S3).

We plotted in Fig. 2 the fetal DNA fraction calculated from the overrepresentation of trisomic chromosome in aneuploid pregnancies and the underrepresentation of chromosome X and the presence of chromosome Y for male pregnancies against gestational age. The average fetal DNA fraction for each sample correlates with gestational age ($P = 0.0051$), a trend that is also previously reported (13).

**Size Distribution of Cell-Free Plasma DNA.** We analyzed the sequencing libraries with a commercial lab-on-a-chip capillary electrophoresis system. There is a striking consistency in the peak fragment size, as well as the distribution around the peak, for all plasma DNA samples, including those from pregnant women and male donor. The peak fragment size was, on average, 261 bp (range: 256–264 bp) (Fig. S4). Subtracting the total length of the Solexa adaptors (92 bp) from 261 bp gives 169 bp as the actual peak fragment size. This size corresponds to the length of DNA wrapped in a chromatosome, which is a nucleosome bound to a H1 histone (24). Because the library preparation includes an 18-cycle PCR, there are concerns that the distribution might be biased. To verify that the size distribution observed in the electropherograms is not an artifact of PCR, we also sequenced cell-free plasma DNA from a pregnant woman carrying a male fetus by using the 454 platform. The sample preparation for this system uses emulsion PCR, which does not require competitive amplification of the sequencing libraries and creates product that is largely independent of the amplification efficiency. The size distribution of the reads mapped to unique locations of the human genome resembled those of the Solexa sequencing libraries, with a predominant peak at 176 bp, after subtracting the length of 454 universal adaptors (Fig. 3 and Fig. S5). These findings suggest that the majority of cell-free DNA in the plasma is derived from apoptotic cells, in accordance with previous findings (22, 23, 25, 26).

Of particular interest is the size distribution of maternal and fetal DNA in maternal cell-free plasma. Two groups have previously shown that the majority of fetal DNA has size range of that of mononucleosome (<200–300 bp), whereas maternal DNA is longer (22, 23). Because 454 sequencing has a targeted read length of 250 bp, we interpreted the small peak at ≈250 bp (Fig. 3 and Fig. S5) as the instrumentation limit from sequencing higher-molecular-mass fragments. We plotted the distribution of all reads and those mapped to Y chromosome (Fig. 3). We observed a slight depletion of Y-chromosome reads in the higher end of the distribution. Reads <220 bp constitute 94% of Y-chromosome and 87% of the total reads. Our results are not in complete agreement with previous findings in that we do not see as dramatic an enrichment of fetal DNA at short lengths (22, 23). Future studies will be needed to resolve this point and to eliminate any potential residual bias in the 454 sample preparation process, but it is worth noting that the ability to sequence single plasma samples permits one to measure the distribution in length enrichments across many individual

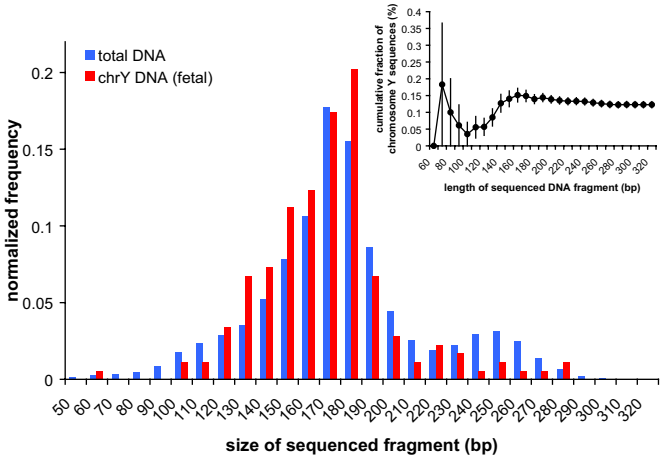


**Fig. 3.** Size distribution of maternal and fetal DNA in maternal plasma. A histogram showing the size distribution of total and chromosome Y-specific fragments obtained from 454 sequencing of maternal plasma DNA from a normal male pregnancy is presented. The distribution is normalized to sum to 1. The numbers of total reads and reads mapped to the Y chromosome are 144,992 and 178, respectively. (*Inset*) Cumulative fetal DNA fraction as a function of sequenced fragment size. The error bars correspond to the standard error of the fraction estimated, assuming that the error of the counts of sequenced fragments follow Poisson statistics.

patients rather than measuring the average length enrichment of pooled patient samples.

**Cell-Free Plasma DNA Shares Features of Nucleosomal DNA.** Because our observations of the size distribution of cell-free plasma DNA suggested that plasma DNA is mainly apoptotic in origin, we investigated whether features of nucleosomal DNA and positioning are found in plasma DNA. One such feature is nucleosome positioning around transcription start sites. Experimental data from yeast and human have suggested that nucleosomes are depleted in promoters upstream of transcription start sites, and nucleosomes are well positioned near transcription start sites (27–30). We applied a 5-bp window spanning ±1,000 bp of transcription start sites of all RefSeq genes and counted the number of tags mapping to the sense and antisense strands within each window. A peak in the sense strand represents the beginning of a nucleosome, whereas a peak in the antisense strand represents the end. After smoothing, we saw that for most plasma DNA samples, at least three well positioned nucleosomes downstream of transcription start sites could be detected, and in some cases, up to five well positioned nucleosomes could be detected, in rough accordance with the results of Schones *et al.* (27) (Fig. 4 and Fig. S6). We applied the same analysis on sequence tags of randomly sheared genomic DNA and observed no obvious pattern in tag localization, although the density of tags was higher at the transcription start site (Fig. 4).

## Discussion

Noninvasive prenatal diagnosis of aneuploidy has been a challenging problem because fetal DNA constitutes a small percentage of total DNA in maternal blood (13), and intact fetal cells are even rarer (6, 7, 9, 31, 32). We showed in this study the successful development of a truly universal, polymorphism-independent noninvasive test for fetal aneuploidy. By directly sequencing maternal plasma DNA, we could detect fetal trisomy 21 as early as the 14th week of gestation. The use of cell-free DNA instead of intact cells allows one to avoid complexities associated with microchimerism and foreign cells that might have colonized the mother; these cells occur at such low numbers that their contribution to the cell-free DNA is negligible (33, 34). Furthermore, there is evidence that cell-free fetal DNA clears from the blood to undetectable levels within a few hours of delivery and therefore is not carried forward from one pregnancy to the next (35–37).

Rare forms of aneuploidy caused by unbalanced translocations and partial duplication of a chromosome are, in principle, detectable by the approach of shotgun sequencing, because the density of sequence tags in the triplicated region of the chromosome would be higher than the rest of the chromosome. Detecting incomplete aneuploidy caused by mosaicism is also possible in principle but may be more challenging, because it depends not only on the concentration of fetal DNA in maternal plasma but also the degree of fetal mosaicism. Further studies are required to determine the effectiveness of shotgun sequencing in detecting these rare forms of aneuploidy.

An advantage of using direct sequencing to measure aneuploidy noninvasively is that it is able to make full use of the sample, whereas PCR-based methods analyze only a few targeted sequences. In this study, we obtained on average 5 million reads per sample in a single run, of which ≈66,000 mapped to chromosome 21. Because those 5 million reads represent only a portion of one human genome, in principle less than one genomic equivalent of DNA is sufficient for the detection of aneuploidy by using direct

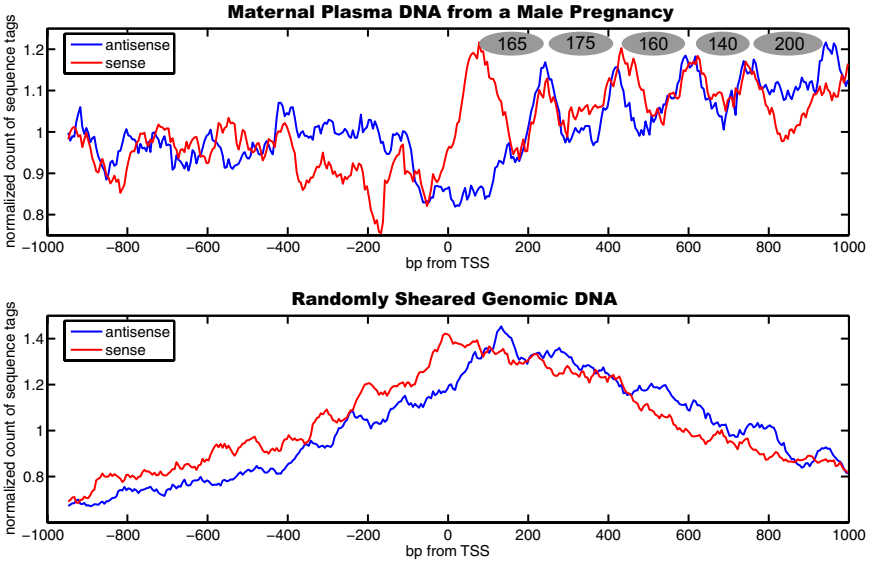


**Fig. 4.** Distribution of sequence tags around transcription start sites (TSS) of ReSeq genes on all autosomes and chromosome X from plasma DNA sample of a normal male pregnancy (*Upper*) and randomly sheared genomic DNA control (*Lower*). The number of tags within each 5-bp window was counted within ±1,000-bp region around each TSS, taking into account the strand to which each sequence tag mapped. The counts from all transcription start sites for each 5-bp window were summed and normalized to the median count among the 400 windows. A moving average was used to smooth the data. A peak in the sense strand represents the beginning of a nucleosome, whereas a peak in the antisense strand represents the end of a nucleosome. In the plasma DNA sample shown here, five well positioned nucleosomes are observed downstream of transcription start sites and are represented as gray ovals. The number within each oval represents the distance in base pairs between adjacent peaks in the sense and antisense strands, corresponding to the size of the inferred nucleosome. No obvious pattern is observed for the genomic DNA control.

sequencing. In practice, a larger amount of DNA was used because there is sample loss during sequencing library preparation, but it may be possible to further reduce the amount of blood required for analysis.

We observed that certain chromosomes have large variations in the counts of sequenced fragments from sample to sample, and that this depends strongly on the GC content (Fig. S1 *A–C*). It is unclear at this point whether this stems from PCR artifacts during sequencing library preparation or cluster generation or the sequencing process itself or whether it is a true biological effect relating to chromatin structure. We strongly suspect that it is an artifact because we also observe GC bias on genomic DNA control, and such bias on the Solexa sequencing platform has recently been reported (38, 39). It has a practical consequence because the sensitivity to aneuploidy detection will vary from chromosome to chromosome; fortunately the most common human aneuploidies (such as 13, 18, and 21) have low variation and therefore high detection sensitivity. Both this problem and the sample-volume limitations may possibly be resolved by the use of single-molecule sequencing technologies, which do not require the use of PCR for library preparation (40).

Plasma DNA samples used in this study were obtained ≈15–30 min after amniocentesis or chorionic villus sampling. Because these invasive procedures disrupt the interface between the placenta and maternal circulation, there have been discussions whether the amount of fetal DNA in maternal blood might increase after invasive procedures. Neither of the studies to date have observed a significant effect (41, 42). Our results support this conclusion, because using the digital PCR assay, we estimated that fetal DNA constituted ≤10% of total cell-free DNA in the majority of our maternal plasma samples. This is within the range of previously reported values in maternal plasma samples obtained before invasive procedures (13). It would be valuable to have a direct measurement addressing this point in a future study.

The average fetal DNA fraction estimated from sequencing data of sex chromosomes are higher than the values estimated from digital PCR data by an average factor of two ($P < 0.005$, paired $t$ test on all male pregnancies that have complete set of data). One possible explanation for this is that the PCR step during Solexa library preparation preferentially amplifies shorter fragments, which others have found to be enriched for fetal DNA (22, 23). Our own measurements of length distribution on one sample do not support this explanation, but we also cannot reject it at this point. It should also be pointed out that using the sequence tags, we find some variation of fetal fraction even in the same sample depending on which chromosome we use to make the calculation (Fig. 2, Fig. S3 and Table S1). This is most likely because of artifacts and errors in the sequencing and mapping processes, which are substantial—recall that only half of the sequence tags map to the human genome with one error or less. Finally, it is also possible that the PCR measurements are biased because they are only sampling a tiny fraction of the fetal genome. These discrepancies will be sorted out in future studies as sequencing reliability improves, and our results show that they do not materially affect the ability to determine fetal aneuploidy.

Our sequencing data suggest that the majority of cell-free plasma DNA is of apoptotic origin and shares features of nucleosomal DNA. Because nucleosome occupancy throughout the eukaryotic genome is not necessarily uniform and depends on factors such as function, expression, or sequence of the region (30, 43), the representation of sequences from different loci in cell-free maternal plasma may not be equal, as one usually expects in genomic DNA extracted from intact cells. Thus, the quantity of a particular locus may not be representative of the quantity of the entire chromosome, and care must be taken when one designs assays for measuring gene dosage in cell-free maternal plasma DNA that target only a few loci.

Historically, because of risks associated with chorionic villus sampling and amniocentesis, invasive diagnosis of fetal aneuploidy was primarily offered to women who were considered at risk of carrying an aneuploid fetus based on evaluation of risk factors such as maternal age, levels of serum markers, and ultrasonographic findings. Recently, an American College of Obstetricians and Gynecologists Practice Bulletin recommended that "invasive diagnostic testing for aneuploidy should be available to all women, regardless of maternal age" and that "pretest counseling should include a discussion of the risks and benefits of invasive testing compared with screening tests" (2). A noninvasive genetic test based on the results described here and in future large-scale studies would presumably carry the best of both worlds: minimal risk to the fetus while providing true genetic information. The costs of the assay are already fairly low; the sequencing cost per sample is approximately $700, and the cost of sequencing is expected to continue to drop dramatically in the near future.

In conclusion, we demonstrated the use of massively parallel sequencing to detect fetal aneuploidy noninvasively with maternal cell-free plasma DNA. Shotgun sequencing can potentially reveal many more previously unknown features of cell-free nucleic acids such as plasma mRNA distributions, as well as epigenetic features of plasma DNA such as DNA methylation and histone modification, in fields including perinatology, oncology, and transplantation, thereby improving our understanding of the basic biology of pregnancy, early human development, and disease.

## Materials and Methods

**Subject Enrollment.** The study was approved by the Institutional Review Board of Stanford University. Pregnant women at risk for fetal aneuploidy were recruited at the Lucile Packard Children's Hospital Perinatal Diagnostic Center of Stanford University during the period of April 2007 to May 2008. Informed consent was obtained from each participant before the blood draw. Blood was collected 15–30 min after amniocentesis or chorionic villus sampling except for one sample that was collected during the third trimester. Karyotype analysis was performed via amniocentesis or chorionic villus sampling to confirm fetal karyotype. Nine T21, 2 T18, 1 T13, and 6 normal singleton pregnancies were included in this study. The gestational age of the subjects at the time of blood draw ranged from 10 to 35 weeks (Table S1). A blood sample from a male donor was obtained from the Stanford Blood Center.

**Sample Processing and DNA Quantification.** Seven to 15 ml of peripheral blood drawn from each subject and donor was collected in EDTA tubes. Blood was centrifuged at 1,600 × *g* for 10 min. Plasma was transferred to microcentrifuge tubes and centrifuged at 16,000 × *g* for 10 min to remove residual cells. The two centrifugation steps were performed within 24 h after blood collection. Cell-free plasma was stored at −80°C until further processing and was frozen and thawed only once before DNA extraction. DNA was extracted from cell-free plasma by using the QIAamp DNA Micro kit (Qiagen) or the NucleoSpin Plasma kit (Macherey–Nagel) according to the manufacturers' instructions. Genomic DNA was extracted from 200 μl of whole blood of the donors by using the QIAamp DNA Blood Mini kit (Qiagen). Microfluidic digital PCR (Fluidigm) was used to quantify the amount of total and fetal DNA by using TaqMan assays targeting at the EIF2C1 locus on chromosome 1 (forward: 5′-GTTCGGCTTTCACCAGTCT-3′; reverse: 5′-CTCCATAGCTCTCCCCACTC-3′; probe: 5′-HEX-CGCCCTGCCATGTGGAAGAT-BHQ1–3′; amplicon size: 81 bp) and the SRY locus on chromosome Y (forward: 5′-CGCTTAACATAGCAGAAGCA-3′; reverse: 5′-AGTTTCGAACTCTGGCACCT-3′; probe: 5′-FAM-TGTCGCACTCTCCTTGTTTTTGACA-BHQ1–3′; amplicon size: 84 bp), respectively. A TaqMan assay targeted at DYS14 (forward: 5′-ATCGTCCATTTCCAGAATCA-3′; reverse: 5′-GTTGACAGCCGTGGAATC-3′; probe: 5′-FAM-TGCCACAGACTGAACTGAATGATTTTC-BHQ1–3′; amplicon size: 84 bp); a multicopy locus on chromosome Y was used for the initial determination of fetal sex from cell-free plasma DNA with traditional real-time PCR. PCR reactions were performed with 1× iQ Supermix (Bio-Rad), 0.1% Tween-20 (microfluidic digital PCR only), 300 nM primers, and 150 nM probes. The PCR thermal cycling protocol was 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Primers and probes were purchased form IDT.

**Sequencing.** A total of 19 cell-free plasma DNA samples, including 18 from pregnant women and 1 from a male blood donor, and genomic DNA sample from whole blood of the same male donor, were sequenced on the Solexa/Illumina platform. ≈1–8 ng of DNA fragments extracted from 1.3–3.2 ml of cell-free plasma was used for sequencing library preparation (Table S1). Library preparation was carried out according to the manufacturer's protocol with slight mod-

ifications. Because cell-free plasma DNA was fragmented in nature, no further fragmentation by nebulization or sonication was done on plasma DNA samples. Genomic DNA from male donor's whole blood was sonicated (Misonix XL-2020) (24 cycles of 30-s sonication and 90-s pause), yielding fragments with size between 50 and 400 bp, with a peak at 150 bp. Approximately 2 ng of the sonicated genomic DNA was used for library preparation. Briefly, DNA samples were blunt-ended and ligated to universal adaptors. The amount of adaptors used for ligation was 500 times less than that written on the manufacturer's protocol. Eighteen cycles of PCR were performed to enrich for fragments with adaptors by using primers complementary to the adaptors. The size distributions of the sequencing libraries were analyzed with a DNA 1000 kit on the 2100 Bioanalyzer (Agilent) and quantified with microfluidic digital PCR (Fluidigm). The libraries were then sequenced by using the Solexa 1G Genome Analyzer according to the manufacturer's instructions.

Cell-free plasma DNA from a pregnant woman carrying a normal male fetus was also sequenced on the 454/Roche platform. Fragments of DNA extracted from 5.6 ml of cell-free plasma (equivalent to ≈4.9 ng of DNA) were used for sequencing library preparation. The sequencing library was prepared according to the manufacturer's protocol, except that no nebulization was performed on the sample, and quantification was done with microfluidic digital PCR instead of capillary electrophoresis. The library was then sequenced on the 454 Genome Sequencer FLX System according to the manufacturer's instructions.

**Data Analysis.** ***Shotgun sequence analysis.*** Solexa sequencing produced 36- to 50-bp reads. The first 25 bp of each read was mapped to the human genome build 36 (hg18) by using ELAND from the Solexa data analysis pipeline. The reads that were uniquely mapped to the human genome having, at most, one mismatch were retained for analysis. To compare the coverage of the different chromosomes, a sliding window of 50 kb was applied across each chromosome, except in regions of assembly gaps and microsatellites, and the number of sequence tags falling within each window was counted, and the median value was chosen to be the representative of the chromosome. Because the total number of sequence tags for each sample was different, for each sample, we normalized the sequence tag density of each chromosome (except chromosome Y) to the median sequence tag density among autosomes. The normalized values were used for comparison among samples in subsequent analysis. We estimated fetal DNA fraction from chromosome 21 for T21 cases, chromosome 18 from T18 cases, chromosome 13 from T13 case, and chromosomes X and Y for male pregnancies. For chromosome 21,18, and 13, the fetal DNA fraction was estimated as $2 \times (x - 1)$, where $x$ was the ratio of the overrepresented chromosome sequence tag density of each trisomy case to the median chromosome sequence tag density of the all disomy cases. For chromosome X, fetal DNA was estimated as $2 \times$ (1-x), where x was the ratio of chromosome X sequence tag density of each male pregnancy to the median chromosome X sequence tag density of all female pregnancies. For chromosome Y, fetal DNA fraction was estimated as the ratio of chromosome Y sequence tag density of each male pregnancy to that of male donor plasma DNA. Because a small number of chromosome Y sequences were detected in female pregnancies, we considered only sequence tags falling within transcribed regions on chromosome Y and subtracted the median number of tags in female pregnancies from all samples; this amounted to a correction of a few percent. The width of 99% confidence intervals was calculated for all disomy 21 pregnancies as $t \times s/\sqrt{N}$, where $N$ is the number of disomy 21 pregnancies, $t$ is the $t$ statistic corresponding to $\alpha = 0.005$ with degree of freedom equal to $N - 1$, and $s$ is the standard deviation. To investigate the distribution of sequence tags around transcription start sites, a sliding window of 5 bp was applied from −1,000 bp to +1,000 bp of transcription start sites of all RefSeq genes on all chromosomes except chromosome Y. The number of sequence tags mapped to the sense and antisense strands within each window was counted. A moving average with a window of 10 data points was used to smooth the data. All analyses were done with Matlab.

For the 454/Roche data, reads were aligned to the human genome build 36 (hg18) by using the 454 Reference Mapper. Reads having an accuracy of ≥90% and coverage (i.e., fraction of read mapped) ≥90% were retained for analysis. To study the size distribution of total and fetal DNA, the number of retained reads falling within each 10-bp window between 50 and 330 bp was counted.

***Genome data retrieval.*** Information regarding GC content, location of transcription start sites of RefSeq genes, location of assembly gaps, and microsatellites were obtained from the University of California, Santa Cruz (UCSC) Genome Browser.

**ACKNOWLEDGMENTS.** We thank the Division of Perinatal Genetics and the General Clinical Research Center of Stanford University for patient recruitment and enrollment; Joseph Derisi, Arend Sidow, Clement Chu, Guang Shi, Norma Neff, and Richard White III for assistance in carrying out sequencing runs; and Kael Fisher and Nick Ingolia for assistance in Solexa pipeline analysis. This work was supported by the Wallace H. Coulter Foundation and the National Institute of Health Director's Pioneer Award. H.C.F. was supported by a Stanford Graduate Fellowship.

1. Cunningham F, *et al.* (2002) in *Williams Obstetrics* (McGraw–Hill, New York), p 942.
2. ACOG Practice Bulletin No. 88 (2007) Invasive prenatal testing for aneuploidy. *Obstet Gynecol* 110:1459–1467.
3. Wapner R, *et al.* (2003) First-trimester screening for trisomies 21 and 18. *N Engl J Med* 349:1405–1413.
4. Alfirevic Z, Neilson JP (2004) Antenatal screening for Down's syndrome. *Br Med J* 329:811–812.
5. Malone FD, *et al.* (2005) First-trimester or second-trimester screening, or both, for Down's syndrome. *N Engl J Med* 353:2001–2011.
6. Herzenberg LA, *et al.* (1979) Fetal cells in the blood of pregnant women: Detection and enrichment by fluorescence-activated cell sorting. *Proc Natl Acad Sci USA* 76:1453–1455.
7. Bianchi DW, *et al.* (1990) Isolation of fetal DNA from nucleated erythrocytes in maternal blood. *Proc Natl Acad Sci USA* 87:3279–3283.
8. Cheung MC, Goldberg JD, Kan YW (1996) Prenatal diagnosis of sickle cell anaemia and thalassaemia by analysis of fetal cells in maternal blood. *Nat Genet* 14:264–268.
9. Bianchi DW, *et al.* (1997) PCR quantitation of fetal cells in maternal blood in normal and aneuploid pregnancies. *Am J Hum Genet* 61:822–829.
10. Bianchi DW, *et al.* (2002) Fetal gender and aneuploidy detection using fetal cells in maternal blood: Analysis of NIFTY I data. National Institute of Child Health and Development Fetal Cell Isolation Study. *Prenat Diagn* 22:609–615.
11. Lo YM, *et al.* (1997) Presence of fetal DNA in maternal plasma and serum. *Lancet* 350:485–487.
12. Dennis Lo YM, Chiu RW (2007) Prenatal diagnosis: progress through plasma nucleic acids. *Nat Rev Genet* 8:71–77.
13. Lo YM, *et al.* (1998) Quantitative analysis of fetal DNA in maternal plasma and serum: Implications for noninvasive prenatal diagnosis. *Am J Hum Genet* 62:768–775.
14. Lo YM, *et al.* (2007) Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med* 13:218–223.
15. Tong YK, *et al.* (2006) Noninvasive prenatal detection of fetal trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: Theoretical and empirical considerations. *Clin Chem* 52:2194–2202.
16. Dhallan R, *et al.* (2007) A non-invasive test for prenatal diagnosis based on fetal DNA present in maternal blood: a preliminary study. *Lancet* 369:474–481.
17. Fan HC, Quake SR (2007) Detection of aneuploidy with digital polymerase chain reaction. *Anal Chem* 79:7576–7579.
18. Lo YM, *et al.* (2007) Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc Natl Acad Sci USA* 104:13116–13121.
19. Quake SR, Fan HC (2006) Non-invasive fetal genetic screening by digital analysis. US Provisional Patent Application No. 60/764:420.
20. Mardis ER (2008) Next-Generation DNA Sequencing Methods. *Annu Rev Genomics Hum Genet* 9:387–402.
21. Lander ES, *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
22. Chan KC, *et al.* (2004) Size distributions of maternal and fetal DNA in maternal plasma. *Clin Chem* 50:88–92.
23. Li Y, *et al.* (2004) Size separation of circulatory DNA in maternal plasma permits ready detection of fetal DNA polymorphisms. *Clin Chem* 50:1002–1011.
24. Cooper G, Hausman R (2007) in *The Cell: A Molecular Approach* (Sinauer, Sunderland, MA), p 168.
25. Jahr S, *et al.* (2001) DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* 61:1659–1665.
26. Giacona MB, *et al.* (1998) Cell-free DNA in human blood plasma: Length measurements in patients with pancreatic cancer and healthy controls. *Pancreas* 17:89–97.
27. Schones DE, *et al.* (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132:887–898.
28. Ozsolak F, Song JS, Liu XS, Fisher DE (2007) High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* 25:244–248.
29. Yuan GC, *et al.* (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309:626–630.
30. Lee W, *et al.* (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* 39:1235–1244.
31. Sohda S, *et al.* (1997) The proportion of fetal nucleated red blood cells in maternal blood: estimation by FACS analysis. *Prenat Diagn* 17:743–752.
32. Hamada H, *et al.* (1993) Fetal nucleated cells in maternal peripheral blood: frequency and relationship to gestational age. *Hum Genet* 91:427–432.
33. Nelson JL (2008) Your cells are my cells. *Sci Am* 298(2):64–71.
34. Khosrotehrani K, Bianchi DW (2003) Fetal cell microchimerism: Helpful or harmful to the parous woman? *Curr Opin Obstet Gynecol* 15:195–199.
35. Lo YM, *et al.* (1999) Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet* 64:218–224.
36. Smid M, *et al.* (2003) No evidence of fetal DNA persistence in maternal plasma after pregnancy. *Hum Genet* 112:617–618.
37. Rijnders RJ, Christiaens GC, Soussan AA, van der Schoot CE (2004) Cell-free fetal DNA is not present in plasma of nonpregnant mothers. *Clin Chem* 50:679–681 and author reply, p 681.
38. Hillier LW, *et al.* (2008) Whole-genome sequencing and variant discovery in *C. elegans*. *Nat Methods* 5:183–188.
39. Dohm JC, Lottaz C, Borodina T, Himmelbauer H (2008) Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* 36(16):e105.
40. Harris TD, *et al.* (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320:106–109.
41. Samura O, *et al.* (2003) Cell-free fetal DNA in maternal circulation after amniocentesis. *Clin Chem* 49:1193–1195.
42. Lo YM, *et al.* (1999) Increased fetal DNA concentrations in the plasma of pregnant women carrying fetuses with trisomy 21. *Clin Chem* 45:1747–1751.
43. Segal E, *et al.* (2006) A genomic code for nucleosome positioning. *Nature* 442:772–778.