Genomic Analysis of Fetal Nucleic Acids in Maternal Blood

Yuk Ming Dennis Lo and Rossa Wai Kwun Chiu

Li Ka Shing Institute of Health Sciences and Department of Chemical Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China; email: loym@cuhk.edu.hk

Keywords

noninvasive prenatal diagnosis, Down syndrome screening, fetal DNA in maternal plasma, massively parallel sequencing

Abstract

The 15 years since the discovery of fetal DNA in maternal plasma have witnessed remarkable developments in noninvasive prenatal diagnosis. An understanding of biological parameters governing this phenomenon, such as the concentration and molecular size of circulating fetal DNA, has guided its diagnostic applications. Early efforts focused on the detection of paternally inherited sequences, which were absent in the maternal genome, in maternal plasma. Recent developments in precise measurement technologies such as digital polymerase chain reaction (PCR) have allowed the detection of minute allelic imbalances in plasma and have catalyzed analysis of single-gene disorders such as the hemoglobinopathies and hemophilia. The advent of massively parallel sequencing has enabled the robust detection of fetal trisomies in maternal plasma. Recent proof-of-concept studies have detected a chromosomal translocation and a microdeletion and have deduced a genome-wide genetic map of a fetus from maternal plasma. Understanding the ethical, legal, and social aspects in light of such rapid developments is thus a priority for future research.
INTRODUCTION

Prenatal diagnosis is an established part of modern obstetrics practice. However, for the direct analysis of fetal DNA, conventional methods require the invasive sampling of fetal tissues using techniques such as amniocentesis or chorionic villus sampling. Such methods are associated with a small but definite risk to the fetus and the mother (33, 70, 87). Such risks have motivated the search for noninvasive prenatal diagnosis methods over the past several decades. Methods such as ultrasound scanning and maternal serum biochemistry are valuable for determining the risks of pregnant women for a number of abnormalities—e.g., fetal trisomy 21 (68)—but cannot allow the direct analysis of fetal DNA. As a result, they are limited in sensitivity and specificity compared with direct methods such as amniocentesis. For a number of years, many groups around the world have searched for fetal nucleated cells in maternal blood. However, the rarity of such cells (8) has severely limited the practicality of this approach. Indeed, in a large multicenter study aimed at the enrichment of such cells for noninvasive prenatal diagnosis, the sensitivity and specificity of this approach were disappointing (7).

In 1997, inspired by the presence of cell-free tumor DNA in the plasma and serum of cancer patients (13, 71), Lo et al. (58) showed that cell-free fetal DNA is present in the plasma and serum of pregnant women. They demonstrated this phenomenon by showing the presence of Y chromosomal DNA sequences in the plasma and serum of women carrying male fetuses. Since then, the discovery of cell-free fetal DNA in maternal plasma has opened up new possibilities for noninvasive prenatal diagnosis. Figure 1 summarizes the time line for key developments in this field.

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>Discovery of circulating fetal DNA (58)</td>
</tr>
<tr>
<td>1998</td>
<td>Measurement of circulating fetal DNA concentrations by real-time PCR (61)</td>
</tr>
<tr>
<td>1999</td>
<td>Development of noninvasive fetal RHD genotyping (34, 59)</td>
</tr>
<tr>
<td>2000</td>
<td>Rapid clearance of circulating fetal DNA following delivery (63)</td>
</tr>
<tr>
<td>2001</td>
<td>Discovery of circulating fetal RNA (80)</td>
</tr>
<tr>
<td>2002</td>
<td>Development of DNA methylation markers for circulating fetal DNA (79)</td>
</tr>
<tr>
<td>2003</td>
<td>Detection of placental mRNA in maternal plasma (73)</td>
</tr>
<tr>
<td>2004</td>
<td>Finding that fetal DNA is shorter than circulating maternal DNA (11, 54)</td>
</tr>
<tr>
<td>2005</td>
<td>Development of hypomethylated SERPINB5 as the first universal fetal DNA marker (16)</td>
</tr>
<tr>
<td>2006</td>
<td>Detection of trisomy 18 by epigenetic allelic ratio method (89)</td>
</tr>
<tr>
<td>2007</td>
<td>Development of trisomy 21 by RNA-SNP method (62)</td>
</tr>
<tr>
<td>2008</td>
<td>Development of massively parallel sequencing for trisomy detection (19, 35)</td>
</tr>
<tr>
<td>2009</td>
<td>Development of relative mutation dosage for monogenic diseases (67)</td>
</tr>
<tr>
<td>2010</td>
<td>Genome-wide identification of fetal DNA methylation markers (76)</td>
</tr>
<tr>
<td>2011</td>
<td>Decoding of fetal genome-wide genetic map from maternal plasma (57)</td>
</tr>
<tr>
<td></td>
<td>Large-scale validation of massively parallel sequencing for aneuploidy detection and introduction as clinical service (18, 32, 51, 75, 84)</td>
</tr>
</tbody>
</table>

BIOLOGY OF FETAL DNA IN MATERNAL PLASMA

Through the use of real-time polymerase chain reaction (PCR) for Y chromosomal sequences, circulating cell-free fetal DNA has been found to be present in maternal plasma at surprisingly high absolute and fractional concentrations, the latter reaching a mean of 3%–6% (61). The recent use of even more precise methods of quantification, such as digital PCR and massively parallel DNA sequencing, has shown the...
fractional fetal concentrations in maternal plasma to be some two- to threefold higher (18, 65, 75). The latter figures are particularly good news for diagnostic applications that require the use of precise quantification—e.g., the detection of fetal chromosomal aneuploidies—and are discussed in detail below. As gestational age progresses, the absolute concentrations of circulating fetal DNA in maternal plasma increase, probably owing to the increased mass of fetal tissues that are releasing DNA into the maternal circulation (3, 61).

Following delivery, fetal DNA in maternal plasma decreases extremely rapidly, with a mean half-life of approximately 16 min (63). Such rapid clearance kinetics suggests that cell-free fetal DNA in maternal plasma would not persist from one pregnancy to the next. This information is important for prenatal diagnostic purposes because possible persistence from prior pregnancies would adversely impact the accuracy of prenatal diagnosis carried out using maternal plasma DNA. Indeed, the lack of persistence of cell-free fetal DNA in maternal plasma represents the majority opinion of groups working in the field (4, 46, 85). There is, nonetheless, one report of the apparent persistence of fetal DNA in maternal plasma following delivery (47). The result from this latter report might be an artifact of the detection of persisting fetal nucleated cells rather than cell-free fetal DNA (50). Another piece of evidence consistent with the lack of persistence of cell-free fetal DNA in maternal plasma represents the majority opinion of groups working in the field (4, 46, 85). There is, nonetheless, one report of the apparent persistence of fetal DNA in maternal plasma following delivery (47). The result from this latter report might be an artifact of the detection of persisting fetal nucleated cells rather than cell-free fetal DNA (50).

The most likely source of circulating fetal DNA in maternal plasma is the placenta. There are a number of lines of evidence that point toward this conclusion. First, circulating fetal DNA molecules have been shown to bear placental epigenetic signatures (10, 14, 16, 20, 74, 77, 91). Second, in anembryonic pregnancies that lack a fetus but possess a placenta, the fetal DNA concentrations in maternal plasma appear to be similar to those in normal pregnancies (2). Third, in cases in which the fetus and placenta exhibit different chromosomal constitutions, e.g., in confined placental mosaicism, the placental genetic signature is consistent with that in the maternal plasma DNA (41, 69).

The source of the maternally derived DNA in maternal plasma during pregnancy has not been directly elucidated. However, in a sex-mismatched bone marrow transplantation model system, it has been shown that the sex genotype of the plasma DNA follows that of the transplantation donor (64). This result shows that the predominant source of plasma DNA in bone marrow transplantation recipients is hematopoietic in origin. At present, with the absence of data to the contrary, this conclusion has been extrapolated to the pregnancy context. The validity of this assumption can be seen indirectly in the search for fetal DNA methylation and RNA markers for detection in maternal plasma, in which investigators have generally used maternal blood cells as the representative maternal tissue for analysis (14, 74, 76, 93).

Cell-free DNA in plasma, either during pregnancy or outside pregnancy, has been shown to consist of short DNA fragments (11, 35, 48, 54, 57, 86). Interestingly, fetal DNA molecules in maternal plasma are shorter than the background maternally derived DNA molecules (11, 54). Paired-end massively parallel sequencing has allowed the analysis of plasma DNA size distribution at single-nucleotide resolution (57). Such analysis has shown that the DNA molecules in maternal plasma (mainly maternally derived) have a size distribution that exhibits a number of size peaks. The most prominent peak is at 166 bp. The next prominent peak occurs at 143 bp, followed by a series of smaller peaks at a periodicity of 10 bp. The main difference between the fetal and maternal DNA molecules is due to a reduction of the 166-bp peak; this peak has been postulated to represent the DNA that is wrapped around a nucleosomal core unit (approximately 146 bp) plus a linker fragment of DNA (20 bp) (57). It has been suggested that
the relative reduction of the 166 bp for fetal DNA compared with maternal DNA might be due to a reduction in the proportion of fetal DNA molecules with the intact linker region. The series of peaks at 10-bp periodicity has been suggested to be related to the nuclease-mediated cleavage of DNA wrapped around the nucleosome core. The conclusion concerning the size-distribution characteristics of the placenta- tionally derived fetal DNA and the presumably hematopoietically derived maternal DNA in plasma has been recently extended to other clinical scenarios. Thus, it has been demonstrated in hematopoietic stem cell transplantation and liver transplantation that nonhematopoietic and hematopoietic DNA molecules in plasma generally take on size-distribution characteristics similar to those of the fetal and maternal DNA, respectively, in maternal plasma (101). With a detailed understanding of the size distributions of nonhematopoietic and hematopoietic DNA in plasma, it might be possible to devise new strategies for the differential enrichment of circulating DNA of different tissue origins from plasma.

DETECTION OF MONOGENIC CHARACTERISTICS

The following sections review the clinical applications for the use of fetal DNA in maternal plasma and serum for the noninvasive analysis of specific fetal genomic features. The applications that would be the easiest to implement are those involving the detection of DNA sequences that the fetus has inherited from its father and that are absent in the maternal genome. The first example of such an application is the determination of fetal sex through the detection of Y chromosomal DNA sequences from maternal plasma and serum (26, 58, 61, 81). The determination of fetal sex is useful for the prenatal investigation of sex-linked diseases in which the detection of a male fetus would indicate a potentially at-risk fetus, whereas a diagnosis that the fetus is female would avoid the need for additional invasive prenatal testing (26). The diagnosis of fetal sex is also useful for the prenatal management of pregnancies at risk for congenital adrenal hyperplasia, in which antenatal dexamethasone treatment could reduce the risk of virilization of a female fetus (81). The diagnostic performance of fetal sexing using cell-free fetal DNA in maternal plasma and serum has recently been studied in a systemic review and meta-analysis for studies published between 1997 and 2011 (28). From the analysis of 57 selected studies, involving 3,524 male-bearing pregnancies and 3,017 female-bearing pregnancies, the sensitivity and specificity for fetal sexing are 95.4% and 98.6%, respectively. The study authors conclude that fetal sex could be determined reliably from 7 weeks onward, with diagnostic performance improving with progress in gestational age. In a number of countries, fetal sexing using maternal plasma DNA analysis has already been available for a number of years. Indeed, preliminary cost analysis has suggested that this method of noninvasive prenatal diagnosis compares favorably with conventional invasive methods in terms of cost-effectiveness, with the added advantage of being safer (45).

Another paternally inherited fetal gene target that has been detected in maternal plasma is the \( RHD \) gene (34, 59). This approach is useful for the detection of an RhD-positive fetus (who has inherited the \( RHD \) gene from its father) carried by an RhD-negative pregnant woman (who typically does not have the \( RHD \) gene). Legler et al. (52) reviewed studies describing the use of this approach for fetal \( RHD \) genotyping from 2006 to 2008 and found that most studies reported a diagnostic accuracy of over 97%. This high diagnostic accuracy has continued with the more recent publications (1, 25, 96). Noninvasive prenatal fetal \( RHD \) genotyping has been clinically used since 2001 (39) and is currently available in a number of centers in Europe and the United States. This technology has also been found to be highly reliable for noninvasive prenatal \( RHCE \) and \( KEL \) genotyping (44, 83).

It is also possible to determine the risk of a fetus being affected by a single-gene disease through maternal plasma DNA analysis. For
autosomal dominant disorders, the approach can be used to detect a mutation that the fetus has inherited from its father but that is absent in the mother’s genome. One example is the detection of the FGFR3 mutation that causes achondroplasia in maternal plasma (56, 82). For autosomal recessive disorders, the detection of a paternally inherited fetal mutation that is absent in the pregnant mother’s genome would indicate that the fetus has a 50% chance of being affected by the disease (9, 22, 31, 53). Invasive prenatal testing would thus be required. A number of different technologies have been used for the detection of a paternally inherited fetal mutation, including restriction enzyme digestion (82), allele-specific base extension followed by mass spectrometry (31), size fractionation–based enrichment of fetal DNA and peptide nucleic acid clamping (53), digital PCR (67), and massively parallel DNA sequencing (57). Conversely, if one could be sure that the paternally inherited fetal mutation is not detectable in maternal plasma despite the presence of detectable levels of circulating fetal DNA, then severe disease in the fetus could be excluded (22).

Ascertaining detectable levels of circulating fetal DNA would require the use of genetic fetal DNA markers [e.g., Y chromosomal markers for a male fetus or single-nucleotide polymorphisms (SNPs)] or DNA methylation markers based on the detection of fetal-specific DNA methylation signatures (10). Apart from detection of the disease-causing mutation, a complementary approach is to detect one or more SNP alleles that are linked to either the mutant or wild-type gene (21, 31, 99).

There are two limitations to the above-mentioned approach to prenatal diagnosis of autosomal recessive diseases through detection of a paternally inherited fetal mutation in maternal plasma: First, it would not provide information on whether the fetus has inherited the maternal mutation, and second, it could not be used in scenarios where both the father and mother carry the same mutation. However, these limitations can be addressed through precise measurement of the relative dosage of the mutant and wild-type alleles in plasma, an approach called the relative mutation dosage (RMD) method (67, 94).

Figure 2a illustrates this approach for an autosomal recessive condition. Let us consider a situation where the father and pregnant mother of a fetus are both heterozygous carriers of the same disease-causing mutation. In the pregnant mother’s plasma, the mutant and wild-type alleles of the disease-causing gene released by her own cells would be present at a ratio of 1:1. Through the liberation of cell-free DNA from the fetus, this allelic ratio may be modified as illustrated in Figure 2a. Thus, if the fetus has inherited the mutant gene from both parents, then the allelic ratio in maternal plasma will be skewed in such a way that the mutant allele would be increased relative to the wild-type allele. Conversely, if the fetus has inherited the wild-type allele from both parents, then the allelic ratio in maternal plasma would show a predominance of the wild-type allele. Finally, if the fetus has inherited one mutant and one wild-type allele from its parents, then the allelic ratio in maternal plasma would remain at 1:1. The extent of allelic skewing positively correlates with the fractional concentration of fetal DNA. This approach could similarly be applied to determine whether the fetus has inherited a mutation from its mother for autosomal dominant conditions. In such a scenario, we would expect the allelic ratio in maternal plasma to be either equal (when the fetus has inherited the maternal mutation) or skewed with more of the wild-type allele (when the fetus has not inherited the mutation).

This strategy can also be applied to sex-linked disorders in which the pregnant mother is a carrier of the disease-causing mutation on one of her X chromosomes (Figure 2b) (94). For such conditions, the at-risk fetus is male and would have only one copy of the X chromosome. Thus, instead of three possibilities regarding the status of allelic skewing in maternal plasma, there would be only two. The first is when the male fetus has inherited the mutant allele from its mother, in which case there would be a predominance of the mutant allele in maternal plasma; the second is when
Noninvasive prenatal diagnosis using the relative mutation dosage (RMD) approach. (a) Noninvasive prenatal diagnosis of autosomal recessive genetic diseases by RMD analysis. The panel shows schematic illustrations of the distributions of the normal/wild-type (N) and mutant (M) alleles in maternal plasma samples containing 10% fetal DNA and obtained from women heterozygous for the mutation and carrying a normal, carrier, or affected fetus. The fetal DNA molecules in each maternal plasma sample are enclosed by the green dashed line. (b) Noninvasive prenatal diagnosis of X-linked diseases by RMD analysis. This panel shows schematic illustrations of the distributions of the N and M alleles in maternal plasma samples containing 10% fetal DNA and obtained from women heterozygous for the mutation and carrying a normal or affected male fetus. The fetal DNA molecules in each maternal plasma sample are enclosed by the green dashed line. (c) Illustrative quantitative parameters for RMD analysis for autosomally inherited and X-linked diseases. Abbreviation: GE, genome equivalent.

To illustrate the calculation, a maternal plasma sample containing a total of 100 GE of DNA with 10% fetal DNA is used.

For the maternal genome, 1 GE includes two individual alleles, i.e., one copy each of the M and N alleles for a heterozygous carrier.

For the fetal genome, 1 GE includes two individual alleles.

For the fetal genome, 1 GE includes 1 X allele, i.e., one copy of either the M or N allele.
the fetus has inherited the wild-type allele from its mother, in which case there would be an overrepresentation of the wild-type allele in maternal plasma. However, for the typical fractional concentrations of fetal DNA in maternal plasma (e.g., 10%), the degree of allelic skewing is low (Figure 2c). Thus, very precise measurement methods are needed to detect such quantitative aberrations. Digital PCR has been used for this purpose (67, 94). Targeted massively parallel sequencing is an alternative technology for such an application, as it has been shown to measure the fractional concentration of circulating fetal DNA precisely (55). RMD analysis is still at a proof-of-concept phase. It is expected that further improvement in technology and the performance of large-scale clinical studies are needed before it can be used as a clinical diagnostic tool.

DETECTION OF CHROMOSOMAL ANEUPLOIDIES

The noninvasive prenatal diagnosis of trisomy 21 has been referred to by many working in the field as the holy grail in prenatal diagnosis. The discovery of fetal DNA in maternal plasma has opened up new possibilities for achieving this goal. However, this is a technologically challenging problem because the diagnosis of trisomy 21 requires the quantitative analysis of fetal chromosome dosage, which is made very difficult by the fact that cell-free fetal DNA represents only a minor fraction of the DNA in maternal plasma (61, 65).

Enrichment of Fetal DNA

To overcome this difficulty, early approaches have focused on the possibility of enriching the fractional concentration of fetal DNA in maternal plasma. Following the finding that fetal DNA in maternal plasma is generally shorter than the circulating maternally derived DNA (11), it was shown that targeting the shorter DNA molecules would allow the enrichment of the fractional fetal DNA concentration (54). However, the degree of enrichment using this approach is relatively modest and has not yet made an impact on solving the problem of detecting fetal trisomy 21 from maternal plasma. Dhallan et al. (29, 30) have proposed that treating maternal blood samples with formaldehyde would reduce the liberation of DNA from maternal blood cells present in the maternal blood sample and would thus lead to a higher fractional concentration of fetal DNA in maternal plasma. However, a number of groups were unable to reproduce this result (17, 24).

DNA Methylation-Based Approaches

Thus, instead of physically enriching the fractional concentration of fetal DNA, several groups have targeted circulating DNA molecules that bear fetal-specific biochemical signatures. In 2002, for example, it was proposed that one approach to selectively target fetal DNA molecules in maternal plasma would be to use differential DNA methylation patterns between the fetal and maternal DNA molecules (79). This approach was demonstrated using a genomic region that exhibited genomic imprinting.

Later, in 2005, it was shown that the promoter of the SERPINB5 gene (coding for maspin) was hypomethylated in the placenta and hypermethylated in maternal blood cells (16). The hypomethylated SERPINB5 promoter region has thus become the first universal marker for fetal DNA in maternal plasma, i.e., one that can be used irrespective of fetal sex and genetic polymorphisms. As the SERPINB5 gene is present on chromosome 18, this marker has led to the detection of trisomy 18 from maternal plasma (89). This report represented the first time that a fetal chromosomal aneuploidy was detected directly from maternal plasma DNA. The trisomy 18 status is determined for a fetus heterozygous for a SNP in the region that exhibits hypomethylation in the placenta and hypermethylation in the maternal blood cells. Thus, a euploid fetus will have an allelic ratio of 1:1, whereas a trisomic fetus will have a ratio of 2:1 or 1:2 (Figure 3). This approach is called the epigenetic allelic ratio (EAR) approach. The limitation of this...
Detection of hypermethylated fetal DNA originating from an aneuploid chromosome

Chromosome dosage assessment

**Figure 3**
Schematic illustrations of epigenetic methods for the noninvasive prenatal detection of fetal chromosomal aneuploidies. The methods rely on the analysis of differentially methylated fetal DNA originating from the potentially aneuploid chromosome; for illustration purposes, a locus that is hypermethylated among the fetal DNA molecules is depicted. The maternal DNA molecules are shown in blue and the fetal DNA molecules are shown in red. The hypermethylated DNA molecules are associated with a filled circle, whereas the hypomethylated DNA molecules are associated with an unfilled circle. In the epigenetic allelic ratio (EAR) approach, the chromosome dosage is assessed by comparing the ratio of heterozygous hypermethylated fetal alleles. In the epigenetic-genetic (EGG) approach, the chromosome dosage is assessed by normalizing the amount of the hypermethylated fetal DNA to a fetal-specific genetic locus, for example, the Y chromosome for male fetuses. For the differentially methylated region (DMR) quantification method, the amount of the hypermethylated fetal DNA molecules in the test sample is compared with that of control samples.

approach is that it can be used only for a fetus that is heterozygous for the SNP, and thus multiple SNPs would be needed to achieve broad population coverage. However, such SNPs are difficult to find because they would need to be present within the genomic region exhibiting differential DNA methylation between the placenta and the pregnant mother’s blood cells.

Since the development of the SERPINB5 marker, many other markers exhibiting differential DNA methylation between the placenta and maternal blood cells have been developed (14, 74, 76, 90, 91). As hypomethylated markers would typically require the use of bisulfite conversion for their detection, with the concomitant risk of DNA degradation (43), there is a recent emphasis on the search for hypermethylated fetal markers that can be detected by relatively simple methylation-sensitive restriction enzyme analysis (10, 90). Despite the large number of such candidate fetal epigenetic markers, the number of markers that have been
validated for detection and fetal-specificity in maternal plasma is relatively limited.

In an effort to overcome the limitations of the EAR approach, the epigenetic-genetic (EGG) chromosome dosage approach has been developed and used to detect trisomies 21 and 18 from maternal plasma (Figure 3) (88, 90, 91). In this approach, a fetal-specific DNA methylation marker on the chromosome involved in the aneuploidy is detected together with a fetal-specific genetic marker in maternal plasma. The latter can be a marker on the Y chromosome (90) or on an autosome (88). Compared with the EAR approach (89), the EGG method has the advantage that the epigenetic and genetic markers do not have to be within the same genomic region. It is thus much easier to address the population coverage issue with the EGG method as there are numerous fetal-specific SNP markers that one could use.

Papageorgiou et al. (77) have recently reported the use of a DNA methylation ratio method for detecting fetal trisomy 21 in maternal plasma (Figure 3). This method uses methylated DNA immunoprecipitation (MeDIP) to precipitate hypermethylated sequences from maternal plasma. Real-time PCR is then used to measure, in the immunoprecipitated DNA, the concentration of sequences on chromosome 21 that have previously been determined to be hypermethylated in the placenta relative to maternal blood cells. The authors reported a remarkably high accuracy for this method in differentiating pregnancies with trisomy 21 fetuses and those with euploid fetuses. It is intriguing that they used maternal whole blood for their analysis instead of plasma; the fractional concentration of fetal DNA in maternal whole blood is much lower than that in maternal plasma. It would also be important to test the reproducibility and suitability of the MeDIP procedure for a high-volume prenatal screening test.

**RNA-Based Approaches**

Fetal mRNA was first reported to be present in maternal plasma in 2000 (80). This was a surprising finding because RNA is generally less stable than DNA, and it was widely expected that any cell-free RNA in plasma would rapidly degrade. However, it has been subsequently shown that plasma RNA is particle-associated and thus might be protected from degradation (72). With the demonstration that the placenta is an important source of circulating mRNA in the plasma of pregnant women (73), a systematic approach for the development of plasma mRNA for noninvasive prenatal diagnosis can be developed using microarray-based expression profiling of the placenta (93). Through the use of such an approach, placenta-specific mRNA transcribed from the PLAC4 gene on chromosome 21 has been detected in maternal plasma (62, 92). With the use of allelic ratio analysis of a SNP on the PLAC4 mRNA, fetal trisomy 21 has been detected in maternal plasma.

This approach, referred to as the RNA-SNP allelic ratio approach, is analogous to the EAR approach described above. As with the EAR approach, its main disadvantage is the need to use multiple SNP markers for broad population coverage, which has proven to be a challenge because of the difficulty of finding a sufficient number of such placenta-specific mRNA markers encompassing SNPs of high enough heterozygosity rates. Nonetheless, the RNA-SNP approach represents the first method that has achieved the direct noninvasive prenatal detection of fetal trisomy 21 from maternal plasma (62). The RNA-SNP concept has also been shown to be feasible for the detection of trisomy 18 (95). Apart from circulating mRNA in maternal plasma, recent data have also indicated the presence of placenta-derived microRNA (miRNA) in maternal plasma (15). Currently, it is not clear whether such circulating miRNA molecules could be used for the prenatal detection of fetal chromosomal aneuploidies.

**Single-Molecule Counting Approaches: Digital Polymerase Chain Reaction and Massively Parallel Sequencing**

For a pregnant woman carrying a fetus with trisomy 21, there would be a slight increase in
the amount of DNA derived from chromosome 21 in maternal plasma compared with that from other chromosomes. The amount of increase is governed by the fractional concentration of fetal DNA. For example, for a maternal plasma sample containing a fractional fetal DNA concentration of 10%, the presence of a trisomy 21 fetus would lead to a 5% increase in the copy number of chromosome 21 in maternal plasma (36, 60). One approach that can detect such a small increase is digital PCR, in which a DNA sample is analyzed by a series of PCRs such that each reaction would primarily contain either one target DNA molecule or no target DNA molecule (97). The number of target DNA molecules in the original DNA sample is then measured by counting the number of positive PCRs. In this type of analysis, the precision of the method is governed by the number of reactions. It has been estimated that 7,680 digital PCRs would allow a DNA sample containing 25% trisomic fetal DNA to be classified correctly 97% of the time (60). The use of digital PCR for the detection of fetal chromosomal aneuploidies has been demonstrated in principle using artificial mixtures of trisomy 21 and euploid DNA (36, 60) and in combination with approaches targeting selected populations of fetal-specific nucleic acid species, including mRNA (60, 92) and DNA methylation (88, 90, 91) markers.

One disadvantage of the digital PCR approach is that a plasma DNA molecule can be counted only if it contains the binding sites for both of the digital PCR primers. Owing to the random fragmentation of plasma DNA, a proportion of the plasma DNA molecules—e.g., those that are very short—would not be analyzed by a given digital PCR assay. Furthermore, in a particular DNA sample, only a very small proportion of the DNA molecules that are derived from the potentially aneuploid chromosome and targeted by the PCR primers would be analyzed. The rest of the molecules on that chromosome would not be analyzed and would thus be wasted.

The recent advent of massively parallel sequencing has provided a much more efficient method for analyzing maternal plasma DNA for aneuploidy detection (19, 35). This approach entails the random sequencing of plasma DNA molecules using massively parallel sequencing, followed by alignment of the sequences to the human genome to identify each sequence’s chromosome of origin (Figure 4). The proportional representation of the potentially aneuploid chromosome is then compared with one or more reference chromosomes, or indeed with the entire genome. An increase in the proportional representation of the potentially aneuploid chromosome would indicate the presence of a trisomic fetus. In the first two reports on the use of massively parallel sequencing for the detection of trisomy 21, the sensitivity and specificity were both 100% (19, 35). However, both studies involved relatively small sample sets. Furthermore, whereas Chiu et al. (19) studied samples from both before and after invasive procedures, Fan et al. (35) studied exclusively samples collected following invasive procedures. In actual clinical prenatal diagnostic scenarios, the samples would not have been subjected to prior invasive procedures; thus, the potential bias introduced by the invasive procedures used by Fan et al. would need to be further evaluated. In any case, large-scale clinical studies are needed to validate these results before massively parallel sequencing of maternal plasma DNA can be used diagnostically.

Over the past three years, a number of studies have been published on the use of massively parallel sequencing of maternal plasma DNA for trisomy detection (Table 1) (12, 18, 23, 32, 51, 75, 84). The data from these studies are very consistent and indicate that as long as the number of aligned sequence reads is 2.3 million or above, the sensitivity and specificity for trisomy 21 detection would be ≥98.6% and ≥97.9%, respectively. As a result of these data, the non-invasive prenatal detection of trisomy 21 using massively parallel sequencing in maternal plasma became clinically available in the United States toward the end of 2011.

The detection of trisomies 18 and 13 using massively parallel sequencing from maternal plasma appears to be more challenging than
Figure 4
Schematic illustration of the procedural framework for using massively parallel genomic sequencing for the noninvasive prenatal detection of fetal chromosomal aneuploidy. (a) Fetal DNA (thick red fragments) circulates in maternal plasma as a minor population in a high background of maternal DNA (black fragments). A sample containing a representative profile of DNA molecules in maternal plasma is obtained. (b) As an illustration, one end of each plasma DNA molecule was sequenced for 36 bp using massively parallel sequencing. The chromosomal origin of each 36-bp sequence was identified through mapping to the human reference genome by bioinformatics analysis. (c) The number of unique sequences mapped to each chromosome was counted and then (d) expressed as a percentage of all unique sequences generated for the sample, termed % chrN for chromosome N. (e) Z-scores for each chromosome and each test sample were calculated using the formula shown. The z-score of a potentially aneuploid chromosome is expected to be higher for pregnancies with an aneuploid fetus (cases E–H, green bars) than those without an aneuploid fetus (cases A–D, blue bars). Figure adapted from Reference 19.
Table 1  Summary of studies investigating the diagnostic performance of massively parallel sequencing of maternal plasma DNA for noninvasive prenatal diagnosis of fetal chromosomal aneuploidies

<table>
<thead>
<tr>
<th>Authors</th>
<th>Analyzer</th>
<th>Protocol(^a)</th>
<th>Mean reads per sample (millions)(^b)</th>
<th>Number of cases</th>
<th>Gestational age (weeks)(^c)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiu et al.</td>
<td>GAIIx</td>
<td>1-plex</td>
<td>2.5</td>
<td>14 T21</td>
<td>14</td>
<td>100</td>
<td>100</td>
<td>19</td>
</tr>
<tr>
<td>Fan et al.</td>
<td>GAIIx</td>
<td>1-plex</td>
<td>5</td>
<td>9 T21</td>
<td>6</td>
<td>100</td>
<td>100</td>
<td>35</td>
</tr>
<tr>
<td>Chiu et al.</td>
<td>SOLiD 3</td>
<td>1-plex</td>
<td>12</td>
<td>5 T21</td>
<td>10</td>
<td>100</td>
<td>100</td>
<td>23</td>
</tr>
<tr>
<td>Lun et al.</td>
<td>GAIIx</td>
<td>1-plex</td>
<td>2.3</td>
<td>1 Rob.</td>
<td>7</td>
<td>100</td>
<td>100</td>
<td>66</td>
</tr>
<tr>
<td>Chiu et al.</td>
<td>GAIIx</td>
<td>2-plex</td>
<td>2.3</td>
<td>86 T21</td>
<td>146</td>
<td>100</td>
<td>97.9</td>
<td>18</td>
</tr>
<tr>
<td>Ehrlich et al.</td>
<td>GAIIx</td>
<td>4-plex</td>
<td>3.5</td>
<td>39 T21</td>
<td>410</td>
<td>100</td>
<td>99.7</td>
<td>32</td>
</tr>
<tr>
<td>Sehnert et al.</td>
<td>GAIIx</td>
<td>1-plex</td>
<td>13–26 (range)</td>
<td>13 T21</td>
<td>25</td>
<td>100</td>
<td>100</td>
<td>84</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>GAIIx</td>
<td>2-plex</td>
<td>4.6</td>
<td>37 T18</td>
<td>227</td>
<td>100</td>
<td>91.9</td>
<td>12</td>
</tr>
<tr>
<td>Peters et al.</td>
<td>HiSeq 2000</td>
<td>1-plex</td>
<td>182(^e)</td>
<td>1 del</td>
<td>7</td>
<td>100</td>
<td>100</td>
<td>78</td>
</tr>
<tr>
<td>Lau et al.</td>
<td>HiSeq 2000</td>
<td>12-plex</td>
<td>2.7</td>
<td>11 T21</td>
<td>76</td>
<td>100</td>
<td>100</td>
<td>51</td>
</tr>
<tr>
<td>Palomaki et al.</td>
<td>HiSeq 2000</td>
<td>4-plex</td>
<td>19</td>
<td>212 T21</td>
<td>1474</td>
<td>98.6</td>
<td>99.8</td>
<td>75</td>
</tr>
</tbody>
</table>

Abbreviations: GAIIx, Genome Analyzer IIx; T21, trisomy 21; T18, trisomy 18; T13, trisomy 13; Rob., Down syndrome due to Robertsonian translocation [(46,XY, der(14;21)(q10;q10), +21)]; 1 del, case with 4.2-Mb deletion on chromosome 12 [del(12)(p11.22p12.1)]; XO, monosomy X.

\(^a\)Multiplexing level refers to the number of samples sequenced jointly in one sequencing lane.

\(^b\)Number of reads used for scoring the disease classification.

\(^c\)Mean/median values shown except when range values are provided as indicated in parentheses.

\(^d\)Scored as “no call” according to protocol.

\(^e\)Information reported only for the affected case.
that of trisomy 21, owing to the susceptibility of many current-generation massively parallel sequencing platforms to the effects of the GC contents of the DNA molecules to be sequenced (19, 23, 35). However, with the appropriate bioinformatics algorithms (12, 37, 84), it appears that trisomies 18 and 13 could be detected with high sensitivity and specificity (Table 1).

The spectrum of fetal chromosomal abnormalities that can be detected from maternal plasma using massively parallel sequencing is likely to increase in the future. In this regard, it has been reported that Down syndrome due to Robertsonian translocation can be detected using this approach (66). It has also been reported that a paternally inherited 4.2-Mb deletion on chromosome 12 could be detected using massively parallel sequencing of maternal plasma DNA (78). However, it is important to note that the blood sample in this case was collected at 35 weeks of gestation and that the case had previously been subjected to amniocentesis at 21 weeks of gestation. Thus, the robustness and general applicability of this approach for the microdeletion syndromes would need to be validated in large-scale studies.

Fetal Genome-Wide Polymorphism and Mutational Profiling

With the rapid reduction in the cost of massively parallel sequencing, it has become increasingly relevant to ask how much of the fetal genome can be elucidated by the deep sequencing of DNA in maternal plasma. In 2010, Lo et al. (57) explored this question by sequencing maternal plasma DNA from a pregnant woman to a sequencing depth equivalent to 65-fold coverage of the haploid human genome (Figure 5a). The father and pregnant mother of the fetus were both heterozygous carriers of β-thalassemia. To assist with the interpretation of the plasma DNA sequencing data, the authors obtained a genetic map of the paternal and maternal genomes by microarray-based SNP genotyping of genomic DNA obtained from the father and pregnant mother. The SNP combinations between the paternal and maternal genomes were then classified into different categories (Figure 5b). Category 1 SNPs were those in which the father and mother were homozygous but for a different allele. For such SNPs, the fetus would be heterozygous and the paternally inherited allele could be used as a fetal genetic marker in maternal plasma. The percentage of category 1 SNPs in which a paternally inherited fetal allele could be seen from the maternal plasma DNA sequencing data could be used to estimate the proportion of the fetal genome that had been covered by the sequencing. Category 1 SNPs could also be used to measure the fractional concentration of fetal DNA in maternal plasma. Category 2 SNPs were those in which the father and mother were homozygous but for the same allele. Such SNPs could be used to estimate the sequencing error rate.

Category 3 SNPs were those in which the father was heterozygous and the mother was homozygous; such SNPs could be used to track the paternally derived inheritance of the fetus. Category 4 SNPs were those in which the father was homozygous and the mother was heterozygous; such SNPs could be used to track the maternally derived inheritance of the fetus. The elucidation of the paternally derived inheritance using category 3 SNPs was much simpler than the analysis of the maternally derived inheritance using category 4 SNPs because the former simply required searching for paternally inherited fetal alleles that were absent in the maternal genome in maternal plasma. The latter, however, required using a quantitative approach in which the inheritance of a maternally derived allele by the fetus may result in the overrepresentation of that allele in maternal plasma, similar to the RMD concept described above. However, as it would not be practical to cover a given SNP as many times as would be needed for RMD (typically on the order of 4,000 times), category 4 SNPs were analyzed as a haplotype, a process called relative haplotype dosage (RHDO) analysis (Figure 5c) (57). This requirement that haplotype information be available for the maternal
Genome would be more readily addressed with the development of a number of genome-wide approaches for direct haplotyping (38, 100).

Category 5 SNPs were those in which both the father and the mother were heterozygous. The resolution of the fetal inheritance of such SNPs would require haplotype information on both the paternal and maternal genomes. Such SNPs would be especially useful for the prenatal diagnosis of autosomal recessive disorders in consanguineous marriage as the haplotype structure in the genomic region carrying a disease-causing gene is likely to be similar for the father and the mother, and thus

![Diagram of haplotype inference](image_url)

**a** Perform paternal genotyping and maternal haplotyping

**b**

1. Assign SNP category
2. Sequence maternal plasma DNA, align, and locate SNP alleles
3. Use category 3 SNP for paternal sequence identification
4. Use category 4 SNP for RHDO analysis to determine maternal inheritance

**c**

<table>
<thead>
<tr>
<th>Loci</th>
<th>Hap N</th>
<th>Hap M</th>
<th>SPRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP 1</td>
<td>20</td>
<td>22</td>
<td>20:22</td>
</tr>
<tr>
<td>SNP 2</td>
<td>23</td>
<td>21</td>
<td>43:43</td>
</tr>
<tr>
<td>SNP 3</td>
<td>19</td>
<td>17</td>
<td>62:60</td>
</tr>
<tr>
<td>N / M</td>
<td>25</td>
<td>21</td>
<td>87:81</td>
</tr>
<tr>
<td>SNP 4</td>
<td>24</td>
<td>25</td>
<td>111:106</td>
</tr>
<tr>
<td>SNP 5</td>
<td>28</td>
<td>24</td>
<td>139:130</td>
</tr>
</tbody>
</table>

**Maternal DNA = 90%**

**Fetal DNA = 10%**

- Haplotype block containing normal allele
- Haplotype block containing mutant allele
- Haplotype block inherited from father
heterozygosity in the paternal genome in one SNP would likely be mirrored by heterozygosity in the maternal genome in the same SNP.

Lo et al. (57) demonstrated that the maternal plasma DNA sequencing data could be analyzed to obtain a genome-wide genetic map of the fetus for SNPs in categories 1–4. Category 5 SNPs were not analyzed in the study owing to the unavailability of the paternal haplotype information. The authors also showed that the maternal plasma DNA sequencing data could be used to demonstrate that the fetus was a heterozygous carrier for \( \beta \)-thalassemia, having inherited the \( HBB \) mutation from the father and the wild-type \( HBB \) gene from the mother.

The main disadvantage of this approach is that it is still relatively expensive to perform such an extensive sequencing of maternal plasma DNA. The recent demonstration that targeted massively parallel sequencing using an in-solution hybridization capture system would allow one to capture fetally and maternally derived DNA in a given genomic region with little bias has opened up the possibility of using this approach to target genomic regions involved in genetic diseases common in a particular population (55). This strategy would potentially greatly reduce the cost of implementing the scanning of selected genomic regions for the noninvasive prenatal diagnosis of genetic diseases as a clinical tool.

**ETHICAL, LEGAL, AND SOCIAL ISSUES**

The development of noninvasive prenatal tests based on the analysis of cell-free fetal nucleic acids in maternal plasma has raised many ethical, legal, and social issues (5, 27, 42). One concern is whether the availability of safe and noninvasive prenatal tests would lead to more pregnant women opting for prenatal testing, and as a consequence might increase the number of subsequent abortions. If indeed more pregnant women would choose noninvasive testing, then it would be important to ensure that there would be a sufficient number of trained genetic counselors for provision of such a service. Another concern is the spectrum of conditions that would be tested in this fashion. Currently, the main applications of such noninvasive prenatal tests are for medical conditions that could have severe morbidities, e.g., fetal chromosomal aneuploidies, severe monogenic...
disorders, and blood group incompatibilities. However, there is a worry by some parties that individual providers of such tests might use the technology for nonmedical indications, e.g., for sex selection and paternity testing (98). This is of especial concern given that certain companies have been reported to offer direct-to-consumer testing of fetal sex using such technologies (6, 49). Another issue that would need to be explored is the source of funding for providing such tests, which will vary from country to country. In countries with a predominantly publicly funded health care system, the cost-effectiveness of using nucleic acid-based noninvasive prenatal diagnosis would be subjected to much scrutiny (45). In view of all of these issues, it is important to ensure that research into the ethical, legal, and social issues of noninvasive prenatal diagnosis parallels the rapid technological developments in this field.

CONCLUSION

The past 15 years have witnessed an extraordinary pace of development for nucleic acid–based noninvasive prenatal diagnosis. Biologically, the fundamental parameters concerning the presence of cell-free fetal nucleic acids in maternal plasma—e.g., molecular size, gestational variation of concentrations, and clearance following delivery—have been elucidated. Diagnostically, this phenomenon has been demonstrated to be applicable to prenatal testing for selected monogenic traits, fetal chromosomal aneuploidies, and even fetal genome-wide genetic and mutational profiling. Remarkably, a number of these applications—notably fetal sexing for sex-linked disorders and congenital adrenal hyperplasia, fetal RHD genotyping, and trisomy 21 detection—have been validated and used clinically within a short span of time. It is thus expected that with time, this technology will play an increasingly important role in prenatal testing, making such testing safer for the fetuses and less stressful for the pregnant women and their families. Future research would further expand the spectrum of disorders that can be detected or monitored using this technology. In particular, the development of new markers for two common pregnancy-associated disorders, preeclampsia and preterm labor, would be a research priority. An intriguing area to explore will be the possible biological functions of circulating fetal nucleic acids in maternal plasma; the presence of circulating placenta-derived miRNA in maternal plasma would be particularly interesting when investigating the potential functionality of such molecules. Finally, it is important that our understanding of the ethical, legal, and social aspects of this field catch up with the technological and scientific developments.

SUMMARY POINTS

1. Cell-free fetal DNA and RNA molecules are present in the plasma of pregnant women.
2. Circulating fetal DNA molecules consist of short DNA fragments derived from the nucleosome.
3. Paternally inherited fetal DNA sequences can be detected in maternal plasma for fetal sexing and fetal RHD genotyping.
4. Prenatal diagnosis of fetal monogenic diseases can be made by precise allelic ratio measurement of the mutant and wild-type alleles using molecular counting technologies such as digital PCR.
5. Massively parallel sequencing of maternal plasma DNA allows the robust detection of fetal trisomies 21, 18, and 13, resulting in the clinical availability of this technology for detection of fetal trisomy 21 in late 2011.
FUTURE ISSUES
1. Research into the ethical, legal, and social issues regarding noninvasive prenatal diagnosis using fetal nucleic acids in maternal plasma will be a priority in the coming years.
2. Research into the use of targeted massively parallel sequencing on maternal plasma DNA would reduce the cost of using molecular counting technology for the prenatal screening of fetuses for selected genetic diseases and chromosomal aberrations.
3. Development of plasma nucleic acid–based markers for detecting common pregnancy-associated disorders such as preeclampsia and preterm labor should be explored.
4. The possible biological functions of circulating fetal nucleic acids in maternal plasma should be explored.

DISCLOSURE STATEMENT
The authors hold patents or have filed patent applications on noninvasive prenatal diagnosis using fetal nucleic acids in maternal plasma. Part of this intellectual property portfolio has been licensed to Sequenom. The authors have received research support from, are consultants to, and hold equities in Sequenom.

ACKNOWLEDGMENTS
The authors are supported by the University Grants Committee of the Government of the Hong Kong Special Administrative Region, China, under the Areas of Excellence Scheme (AoE/M-04/06). Y.M.D.L. is supported by an endowed chair from the Li Ka Shing Foundation.

LITERATURE CITED
37. Fan HC, Quake SR. 2010. Sensitivity of noninvasive prenatal detection of fetal aneuploidy from maternal plasma using shotgun sequencing is limited only by counting statistics. *PLoS ONE* 5:e10439


Contents

Human Genetic Individuality
   Maynard V. Olson ............................................................ 1

Characterization of Enhancer Function from Genome-Wide Analyses
   Glenn A. Mastro, Stephen G. Landt, Michael Snyder, and Michael R. Green .......... 29

Methods for Identifying Higher-Order Chromatin Structure
   Samin A. Sajan and R. David Hawkins ...................................... 59

Genomics and Genetics of Human and Primate Y Chromosomes
   Jennifer F. Hughes and Steve Rozen ...................................... 83

Evolution of the Egg: New Findings and Challenges
   Katrina G. Claw and Willie J. Swanson .................................. 109

Evolution of the Immune System in the Lower Vertebrates
   Thomas Boehm, Norimasa Iwanami, and Isabell Hess .......................... 127

The Human Microbiome: Our Second Genome
   Elizabeth A. Grice and Julia A. Segre .................................... 151

Functional Genomic Studies: Insights into the Pathogenesis
   of Liver Cancer
   Ze-Guang Han ........................................................................ 171

A Comparative Genomics Approach to Understanding Transmissible
   Cancer in Tasmanian Devils
   Janine E. Deakin and Katherine Belov ....................................... 207

The Genetics of Sudden Cardiac Death
   Dan E. Arking and Nona Sotoodehnia ...................................... 223

The Genetics of Substance Dependence
   Jen-Chyong Wang, Manav Kapoor, and Alison M. Goate ...................... 241

The Evolution of Human Genetic Studies of Cleft Lip and Cleft Palate
   Mary L. Marazita .................................................................. 263

Genomic Analysis of Fetal Nucleic Acids in Maternal Blood
   Yuk Ming Dennis Lo and Rosa Wai Kwun Chiu .............................. 285
Enzyme Replacement Therapy for Lysosomal Diseases: Lessons from
20 Years of Experience and Remaining Challenges
R.J. Desnick and E.H. Schuchman ......................................................... 307

Population Identification Using Genetic Data
Daniel John Lawson and Daniel Falush ................................................. 337

Evolution-Centered Teaching of Biology
Karen Burke da Silva .............................................................................. 363

Ethical Issues with Newborn Screening in the Genomics Era
Beth A. Tarini and Aaron J. Goldenberg .............................................. 381

Sampling Populations of Humans Across the World: ELSI Issues
Bartha Maria Knoppers, Ma’n H. Zarwati, and Emily S. Kirby .......... 395

The Tension Between Data Sharing and the Protection
of Privacy in Genomics Research
Jane Kaye .................................................................................................. 415

Genetic Discrimination: International Perspectives
M. Otlowski, S. Taylor, and Y. Bombard .............................................. 433

Indexes

Cumulative Index of Contributing Authors, Volumes 4–13 ................. 455
Cumulative Index of Chapter Titles, Volumes 4–13 ................................. 459

Errata

An online log of corrections to Annual Review of Genomics and Human Genetics articles
may be found at http://genom.annualreviews.org/errata.shtml