Mini Review

Rossa Wai Kwun Chiu* and Yuk Ming Dennis Lo

Clinical applications of maternal plasma fetal DNA analysis: translating the fruits of 15 years of research

Abstract

The collection of fetal genetic materials is required for the prenatal diagnosis of fetal genetic diseases. The conventional methods for sampling fetal genetic materials, such as amniocentesis and chorionic villus sampling, are invasive in nature and are associated with a risk of fetal miscarriage. For decades, scientists had been pursuing studies with goals to develop non-invasive methods for prenatal diagnosis. In 1997, the existence of fetal derived cell-free DNA molecules in plasma of pregnant women was first demonstrated. This finding provided a new source of fetal genetic material that could be obtained safely through the collection of a maternal blood sample and provided a new avenue for the development of non-invasive prenatal diagnostic tests. Now 15 years later, the diagnostic potential of circulating fetal DNA analysis has been realized. Fruitful research efforts have resulted in the clinical implementation of a number of non-invasive prenatal tests based on maternal plasma DNA analysis and included tests for fetal sex assessment, fetal rhesus D blood group genotyping and fetal chromosomal aneuploidy detection. Most recently, research groups have succeeded in decoding the entire fetal genome from maternal plasma DNA analysis which paved the way for the achievement of non-invasive prenatal diagnosis of many single gene diseases. A paradigm shift in the practice of prenatal diagnosis has begun.

Keywords: cell-free fetal DNA; Down syndrome; maternal plasma DNA; next-generation sequencing; non-invasive prenatal diagnosis.

Prenatal diagnosis

To diagnose genetic diseases of an unborn child, one must obtain genetic materials (DNA/RNA) from the child for analysis. The most direct way of obtaining fetal genetic materials is to obtain fetal tissues either through amniocentesis or chorionic villus sampling. While these methods have become established conventional procedures to conduct prenatal diagnosis, by virtue of their invasive nature, they are associated with a risk of fetal miscarriage [1]. To mitigate the risks, researchers have developed screening tests, involving fetal ultrasonography or maternal serum biochemical assessment, to stratify women into groups according to their risks for fetal genetic or chromosomal abnormalities [2]. Invasive prenatal diagnosis is generally recommended as an option for pregnancies identified by the screening tests, or personal and family history, to be in the high-risk groups. Though these screening programs, such as those for Down syndrome, have been successfully implemented in many parts of the world, they are associated with a 3%–5% false-positive rate, i.e., women with healthy fetuses but are labeled as belonging to the high-risk group [2]. Consequently, many women are unnecessarily posed with the question of whether to proceed to invasive testing or not. Other researchers have focused on ways to obtain fetal genetic materials non-invasively. Since the 1970s, efforts have been made to isolate intact fetal cells from maternal blood. However, the number of fetal cells that have entered into the maternal circulation was proven to be too few to allow reliable diagnosis using available cell isolation technologies [3].
Fifteen years ago

In 1997, instead of looking for fetal cells, Lo et al. turned to the acellular component of blood, namely plasma and serum, to look for fetal genetic material [4]. Through the detection of DNA sequences of chromosome Y in plasma of women pregnant with male fetuses, the investigators demonstrated the existence of cell-free DNA originating from the fetus in the maternal circulation. Because the concentrations of such cell-free fetal DNA molecules were subsequently shown to amount to some 10% of the total DNA in maternal plasma [5, 6], the scene for using cell-free fetal DNA analysis to achieve non-invasive prenatal diagnosis was set.

The first step towards achieving non-invasive prenatal diagnosis through circulating fetal DNA analysis was to understand this new source of fetal genetic material. It was learned that fetal DNA molecules were detectable in maternal plasma from about seventh week of gestation and the absolute concentration increases with the advancement of gestation [5, 7]. There is rapid turnover of fetal DNA molecules and they are removed from maternal circulation within hours of delivery [8]. The fetal DNA molecules that circulate in maternal plasma are predominantly derived from the placenta [9, 10]. They circulate in the form of short DNA fragments of <200 bp in length [11]. Recent data suggest that a proportion of fetal DNA molecules are bound to histone proteins in the form of mononucleosomes [12]. These data further suggest that fetal DNA enters into the maternal circulation as a result of placental cell turnover [13]. Consequently, pregnancy-associated complications that involve increased turnover of placenta cells, such as preeclampsia, invasive placenta, are associated with elevated levels of circulating fetal DNA [14].

Detection of paternally inherited traits

Fetal sex determination

While fetal DNA molecules are proven to be present in maternal plasma, they coexist with a large background of maternal DNA predominantly derived from the mother’s blood cells [15]. Thus, one way to reliably detect fetal DNA amidst the many maternal DNA molecules in maternal plasma is to analyze gene sequences that are unique to the fetus and distinguishable from the maternal DNA, i.e., genes that the fetus has inherited from its father. For example, the positive detection of DNA sequences from chromosome Y in a maternal plasma sample suggests the presence of a male fetus [4]. The analysis of chromosome Y sequences in maternal plasma for fetal sex determination has been performed by many groups [16]. A recent meta-analysis of 57 studies confirmed the high accuracy of non-invasive fetal sex determination [7]. The report concluded that fetal sex assessment is less reliable when performed before the seventh week of gestation. The use of real-time quantitative polymerase chain reaction achieved better accuracy than conventional polymerase chain reaction. There were no significant differences between the protocols that detected different sequences on chromosome Y.

Prenatal sex determination is clinically useful in the management of sex-linked disorders, such as hemophilia [17, 18]. Sex-linked disorders result in manifestations mainly in males. If circulating fetal DNA analysis excludes a fetus as being male, an invasive prenatal diagnostic procedure could be avoided. In addition, non-invasive fetal sex determination has been found to be useful in the prenatal management of congenital adrenal hyperplasia [18]. 21-Hydroxylase deficiency is the commonest form of congenital adrenal hyperplasia. One of the manifestations of 21-hydroxylase deficiency is the overproduction of androgens, resulting in virilization of female fetuses which can be avoided with the administration of steroids during early pregnancy. However, steroid administration is not necessary for male fetuses. Thus, knowing the sex of the fetus early in the course of the pregnancy allows clinical decisions to be made regarding the necessity for steroid treatment. Prenatal fetal sex assessment would also facilitate the management of fetuses shown to have ambiguous genitalia on ultrasound. For more than 10 years, non-invasive fetal sex assessment has been clinically applied in many centers around the world [16, 17].

Fetal rhesus D status determination

Another paternally inherited trait that has been assessed routinely through circulating fetal DNA analysis involves the rhesus D blood group [19, 20]. The rhesus D antigen is a protein product of the rhesus D (RHD) gene. A person is rhesus D positive if his/her genome has at least one copy of RHD. A person is rhesus D negative if his/her genome does not have a functional RHD. Rhesus D incompatibility describes the situation when a rhesus D negative woman is pregnant with a rhesus D positive fetus. In a rhesus D negative woman with prior rhesus D antigen sensitization, preformed anti-rhesus D antibodies may cross the placenta and harm the fetus, resulting in hemolytic disease of
the newborn. Because the genome of a rhesus D negative woman does not contain RHD, when RHD sequences are detected in maternal plasma, such sequences are likely to be from the fetus and is inherited from the father. This rationale formed the basis for the development of a circulating fetal DNA test for non-invasive fetal rhesus D status determination.

The clinical use of the non-invasive test for fetal rhesus D status assessment was launched soon after the discovery of circulating fetal DNA [20, 21]. Knowing the fetal rhesus D status influences how a pregnancy could be managed. When a fetus is tested to be rhesus D positive, the pregnancy would need to be closely monitored and anti-D immunoglobulin is typically administered later in pregnancy. On the contrary, when a fetus is tested to be rhesus D negative, the pregnancy may be managed like other uncomplicated pregnancies. Due to the different consequences in terms of management, efforts have been made to improve the reliability of the non-invasive fetal rhesus D test [21, 22]. For example, internal controls to indicate the presence of fetal DNA in maternal plasma samples have been developed to minimize the chance of reporting false-negative results. DNA sequences that have been analyzed as internal controls include chromosome Y sequences for male fetuses [4], panels of polymorphisms to identify alleles not derived from the mother [23, 24] and epigenetic signatures of placental tissues [25]. When a sample is tested negative for RHD, the fetus can be deemed as rhesus D negative only when one is certain that the sample contained fetal DNA by detecting a positive signal in such internal controls. There are rare rhesus D alleles that have a partial sequence or mutation that result in no or weakened expression of rhesus D antigen [21, 22]. The assay for RHD detection may result in positive detection of such alleles while they may not have functional consequences and vice versa. Thus, efforts have been made in developing primer sets that allow the reliable detection of RHD in maternal plasma while offering discrimination of the weak D and partial D alleles for the avoidance of false interpretations of the fetal genotype [21, 22].

The circulating fetal DNA test for fetal rhesus D status determination has been applied for clinical use in many countries [26]. In some centers, the test is performed late in pregnancy, around the 28th week of gestation, to rationalize the administration of the anti-D immunoglobulin only to rhesus D negative women with rhesus D positive fetuses [21]. In other centers, the test is performed early in pregnancy to determine the course of management [22]. Recently, cost-benefit studies have been performed and have concluded that the implementation of the non-invasive fetal rhesus D test is a cost-effective strategy for managing rhesus D incompatibility. National adoption of the test has been recommended in Denmark, Sweden and the Netherlands [21].

### The holy grail of prenatal diagnosis – Down syndrome

Down syndrome detection has been hailed by some as the holy grail of non-invasive prenatal diagnosis [26]. Down syndrome is a common fetal chromosomal aneuploidy affecting 1 in 700 pregnancies on average. It is the commonest reason why couples may be offered the option of invasive prenatal diagnosis. Down syndrome screening programs that involve fetal ultrasound assessment and the measurement of maternal serum biochemical markers have been implemented effectively in many parts of the world [2]. Typically, 3%–5% of all the women screened are identified as high-risk and may warrant an invasive prenatal diagnostic procedure. In other words, many more women are undergoing invasive diagnosis than the incidence of Down syndrome. Thus, researchers have long been trying to develop more effective screening tests to lower the false-positive rate and reduce the number of women needing to undergo invasive prenatal diagnosis unnecessarily.

Down syndrome is typically caused by the presence of a third copy of chromosome 21, hence the name trisomy 21, in genomes of affected individuals. The use of circulating fetal DNA analysis for trisomy 21 detection posed a number of challenges. The fragmented and cell-free nature of circulating fetal DNA prevents the application of techniques involving the counting of intact chromosomes, such as fluorescence in situ hybridization [11]. Furthermore, the coexistence of the high maternal DNA background dilutes the overall increase in chromosome dosage in the sample [27]. Through the years, a number of approaches have been developed to achieve the non-invasive detection of Down syndrome [14]. However, the most promising approach that has made its way to clinical use is based on massively parallel sequencing (MPS), or the so-called next-generation sequencing [28].

MPS is a DNA sequencing technology, i.e., the polynucleotide sequences of DNA molecules are decoded. MPS platforms typically can decode the sequences of short DNA molecules up to hundreds of base pairs but can analyze millions to billions of such DNA molecules in a parallel fashion in each run. Due to the high sequencing capacity of these analyzers, a common application of MPS is whole genome sequencing [28]. However, the analytical
specifications of MPS are in fact well-suited for the analysis of plasma DNA samples [27, 29]. First, plasma DNA molecules are short DNA fragments. Thus, the lengths of plasma DNA molecules are well within the DNA size range suitable for MPS analysis. Second, sequencing allows the identification of the gene or chromosome origin of each sequenced plasma DNA molecule. Third, because many plasma DNA molecules are analyzed during each MPS run, a quantitative profile of the sequenced plasma DNA species could be obtained. Consequently, taking advantage of these analytical features of MPS, methods for trisomy 21 detection based on the sequencing of maternal plasma DNA molecules have been developed [29–31].

The primary goal of the MPS protocols for the non-invasive detection of trisomy 21 is to identify if there is an increase in plasma DNA fragments originating from chromosome 21 in a maternal plasma sample [27]. In general, both the maternal- and fetal-derived DNA molecules in the maternal plasma sample are sequenced. The relative amounts of plasma DNA molecules originating from the various human chromosomes are calculated. In particular, the proportion of DNA molecules originating from chromosome 21 is assessed to determine if there is an increase beyond what is expected for a euploid genome. In most instances, because the mother’s genome is euploid, any deviations in the proportions of maternal plasma DNA molecules originating from any chromosomes would be the result of chromosomal abnormalities in the fetal genome. Thus, after adjustments for the statistical noises inherent to the analytical protocol has been made, observations of increased proportions of maternal plasma DNA molecules originating from chromosome 21 are interpreted as an increase in chromosome 21 dosage in the fetal genome, i.e., a result compatible with fetal trisomy 21.

Down syndrome detection through MPS analysis of maternal plasma DNA can be achieved by randomly sequencing a representative profile of DNA molecules from the maternal plasma sample [29, 32] or by selectively focusing on the sequencing of DNA molecules from several chromosomes only [33, 34]. The latter is also termed as a targeted sequencing approach. The random sequencing protocol is currently most widely practiced and has been shown to be reproducible across different studies by different groups [35–37]. Targeted analysis of selective chromosomes, however, could be achieved either by presequencing amplification of selective chromosomal loci or with the use of hybridization probes designed to capture selective chromosome loci from the plasma DNA pool. In general, targeted sequencing approaches are said to save sequencing costs because the sequencing is performed mainly on the loci of interest [38]. However, efforts and additional reagents are needed for the target selection. Furthermore, the process of target selection may introduce biases between the relative proportions of DNA molecules from the different chromosomes. Such biases would need to be accounted for in interpreting whether the chromosome representation of DNA molecules is abnormal or not. Typically, more vigorous bioinformatics normalization procedures are needed to compensate for such target selection biases [33, 34, 39]. Hence, due to the need for extra experimental and data analysis procedures, coupled with the rapid reduction in the costs of sequencing, it remains to be proven if the targeted MPS protocols are indeed more cost-effective.

In terms of clinical performance, a number of large-scale studies have shown that maternal plasma DNA sequencing for Down syndrome detection is accurate when performed on pregnancies identified by conventional screening modalities to be of high-risk for Down syndrome [35–37]. Combining all the studies published to date, MPS has been applied to analyze plasma collected from hundreds of Down syndrome pregnancies and several thousand normal pregnancies. Maternal plasma DNA sequencing has been shown to be able to detect approximately 99% of the Down syndrome fetuses at a false-positive rate of <1% [40]. The key factors that have been shown to affect the performance of the non-invasive Down syndrome tests are the fractional fetal DNA concentration in the maternal plasma sample and the analytical precision of the test protocol for measuring the proportional representation of chromosome 21 [35]. The chance of false-negative results is higher when the fetal fraction of DNA molecules in maternal plasma is low [35, 36]. Thus, the measurement of the fractional fetal DNA concentration in each tested sample is a useful safeguard to identify cases with particularly low fetal DNA concentrations [36]. MPS protocols that are optimized to have high analytical precision result in better discrimination between the Down syndrome and euploid cases, resulting in higher test accuracy [35].

As a result of the high accuracy shown for the MPS-based non-invasive test for Down syndrome detection, the offering of the test for clinical use has been available since 2011 [36]. To date, the test service is available in the US, parts of Asia and Europe. Based on the clinical evidence published to date and the relative high cost of the test, the test is currently an option for women identified to be at high-risk for Down syndrome based on conventionally accepted screening criteria and tests [35, 36]. In other words, instead of proceeding directly to invasive prenatal diagnosis after having been screened as belonging to the high-risk group, women could pursue the maternal
plasma DNA test. As the test has high specificity [41], the majority of the women carrying a non-affected fetus would be tested negative and may consider not undergoing invasive testing. However, if a woman is tested positive with the maternal plasma DNA test, she has a high likelihood of carrying a fetus with Down syndrome. But because the test is associated with a small false-positive rate, women who are tested positive are advised to pursue confirmatory invasive testing. In the context of the US, according to Palomaki et al. [41], a woman has a seven in eight chance of having an affected fetus if she had been identified as high-risk by the first trimester combined test and further received a positive result by the sequencing test.

Aneuploidies other than trisomy 21

After the non-invasive detection of trisomy 21 has been achieved with the sequencing of maternal plasma DNA, an immediate question raised was whether the same testing algorithm could be applied to the detection of other common aneuploidies, namely Edwards syndrome (trisomy 18) and Patau syndrome (trisomy 13). Chen et al. [42] applied the trisomy 21 sequencing test data analysis protocol onto the analysis of maternal plasma samples collected from women carrying fetuses with trisomy 18 or trisomy 13. The authors reported that the basic algorithm had poor sensitivity for trisomy 18 and trisomy 13. It was then realized that the test performance was affected by a GC bias effect inherent to many MPS platforms [30, 42]. The GC bias effect describes a phenomenon whereby DNA molecules with particularly high or low guanine and cytosine nucleotide contents would result in skewing of their quantitative representation among the sequenced data. The DNA reads originating from chromosomes 18 and 13 are found to suffer from the GC bias effect of MPS and resulted in misrepresentation of their quantitative profile in maternal plasma. This has further led to the incorrect estimation of whether there are increased amounts of such DNA molecules in the sample. To improve the detection of trisomy 18 and trisomy 13 by maternal plasma DNA sequencing, procedures to minimize or to remove the GC bias effect were needed. There are several strategies used to remove the GC bias. For example, the mathematical relationship between the GC content of DNA molecules and the degree of quantitative bias could be determined to develop a formula to bioinformatically correct for the bias [42]. Alternatively, one could selectively sequence or analyze DNA molecules with favorable GC contents [32, 33]. As another option, one could discard the DNA reads originating from molecules with non-favorable GC contents [32]. With the introduction of such GC normalization procedures, the detection of trisomy 18 and trisomy 13 attained reasonably high accuracy [37, 41, 42]. In fact, the GC normalization procedures have been found to further improve the performance of the sequencing test for trisomy 21 also [41].

Besides aneuploidies involving whole chromosomes, researchers have also explored the use of maternal plasma DNA sequencing for the non-invasive detection of aneuploidies involving just a chromosome segment. The protocol used for trisomy 21 detection was found to be useful also for the detection of Down syndrome due to Robertsonian translocations involving just the long arm of chromosome 21 [37, 43]. Peters et al. [44] reported the detection of a significant reduction in sequenced DNA reads from a 4.2 Mb region on chromosome 12 that was deleted in the fetal genome. Jensen et al. [45] also detected a reduction in DNA reads from a 3 Mb region on chromosome 22 within the DiGeorge syndrome locus. However, both studies required a substantially higher number of plasma DNA molecules (20 times that required for trisomy 21 detection) to be analyzed from each maternal plasma sample in order for the change in representation of the small chromosome segment to reach statistical significance.

Single gene diseases

The direct detection of gene mutations transmitted by a father in maternal plasma allows the non-invasive diagnosis of paternally inherited autosomal dominant diseases. The rationale is the same as for the detection of fetal chromosome Y and/or RHD sequences. However, the strategy is insufficient for the non-invasive diagnosis of maternally inherited or autosomal recessive diseases [46]. The challenge to solve is how could one assess the fetal inheritance of maternal alleles despite the presence of the high maternal DNA background in maternal plasma? Lun et al. addressed the challenge with the use of a quantitative approach [47]. The prenatal diagnosis is generally performed in the context of a woman who is a carrier for a mutation associated with a single gene disorder, i.e., she is heterozygous for the mutation. Lun et al. designed digital PCR assays to detect and quantify the number of mutant and normal alleles originating from the disease locus in maternal plasma. The relative amounts between the mutant and normal alleles are compared and hence the approach is termed relative mutation dosage (RMD). If the fetus is also heterozygous for the mutation, it should
contribute additional molecules of the mutant and normal alleles, but in equal amounts, to maternal plasma. Thus, the combined (maternal and fetal) mutant alleles should be of similar quantity as the combined number of normal alleles. On the contrary, if the fetus is homozygous for the normal allele, it would only contribute additional molecules of the normal allele into maternal plasma, resulting in an overrepresentation of the normal allele relative to the mutant allele. Similarly, when the fetus is homozygous for the mutant allele, there will be an overrepresentation of the mutant allele relative to the normal allele. This approach has been shown to be applicable to the non-invasive detection of beta-thalassemia [47], sickle cell anemia [48] and hemophilia [49].

Decoding the fetal genome from maternal plasma

With the successful implementation of maternal plasma DNA sequencing for fetal chromosomal aneuploidy detection and the achievement of the prenatal assessment of maternally inherited alleles, the scene was set to attempt to determine how much more genetic information one could obtain from a maternal plasma DNA sample. Lo et al. [12] used MPS to analyze a maternal plasma sample collected at 12th week of gestation from a pregnant couple where both the woman and her partner were carriers for beta-thalassemia mutations. Billions of plasma DNA molecules were analyzed from the sample. The amount of sequencing performed was equivalent to covering a haploid human genome 65 times. The authors showed that the entire fetal genome was present in maternal plasma at a constant relative proportion to the amount of maternal DNA molecules. Determination of the fetal inheritance of the parental beta-thalassemia mutations was also achieved. However, instead of using the RMD approach to determine the fetal inheritance of the maternal mutation, a method termed relative haplotype dosage (RHDO) was used [12]. RMD requires the analysis of thousands of mutant and normal alleles at the disease locus so that a statistically meaningful comparison between the amounts of the mutant and normal alleles could be made. RHDO, however, reaches a statistically meaningful comparison between the mutant and normal alleles by combining the amounts of sequenced reads obtained from any DNA molecules originating from the haplotype containing the mutant locus or the haplotype containing the normal locus. In other words, even DNA molecules not actually covering the mutation site are used in the statistical comparison. As a result, the fetal inheritance of maternal alleles could be determined with less sequencing effort and hence less cost. In fact, the analysis of the fetal genome by MPS of maternal plasma DNA was shown to be feasible in two additional studies [50, 51].

A 15-year journey and beyond

In a 15-year time frame since the first report of the existence of circulating cell-free fetal DNA in maternal plasma [4], we have witnessed the launch of clinical services covering the non-invasive prenatal assessment of fetal sex, fetal rhesus D status and fetal chromosomal aneuploidies and finally when the entire fetal genome is decoded non-invasively. A paradigm shift in the way prenatal diagnosis is practiced has begun. Translation of the fruits of the research efforts on circulating fetal DNA analysis into clinical use has generally taken the path which includes the initial development of novel analysis approaches, then fine-tuning and improvisation of the methods, subsequent demonstration of the clinical efficacy of the test and lastly followed by efforts to minimize costs. It is hoped that this path to translation will continue to deliver a succession of new prenatal tests. While the decoding of the fetal genome from maternal plasma offers the theoretical possibility of non-invasive diagnosis of any fetal genetic diseases from maternal plasma DNA analysis, it certainly unveils many ethical issues and social dilemmas to be explored [52]. Consequently, while it is exciting that the technological developments in non-invasive prenatal diagnosis have attained new heights, it is no less important that the intertwined ethical and social issues be considered and addressed.

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) and funding from the S.K. Yee Foundation. Y.M.D. Lo is supported by an endowed chair from the Li Ka Shing Foundation.

Conflict of interest statement

Authors’ conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article. Research support played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.
**References**