Review

Non-invasive prenatal diagnosis by single molecule counting technologies

Rossa W.K. Chiu[1,2], Charles R. Cantor[3] and Y.M. Dennis Lo[1,2]

[1] Centre for Research into Circulating Fetal Nucleic Acids, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, New Territories, Hong Kong SAR, China
[2] Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, New Territories, Hong Kong SAR, China
[3] Sequenom Inc., 3595 John Hopkins Court, San Diego, CA 92121-1331, USA

Non-invasive prenatal diagnosis of fetal chromosomal aneuploidies and monogenic diseases by analysing fetal DNA present in maternal plasma poses a challenging goal. In particular, the presence of background maternal DNA interferes with the analysis of fetal DNA. Using single molecule counting methods, including digital PCR and massively parallel sequencing, many of the former problems have been solved. Digital mutation dosage assessment can detect the number of mutant alleles a fetus has inherited from its parents for fetal monogenic disease diagnosis, and massively parallel plasma DNA sequencing enables the direct detection of fetal chromosomal aneuploidies from maternal plasma. The analytical power of these methods, namely sensitivity, specificity, accuracy and precision, should catalyse the eventual clinical use of non-invasive prenatal diagnosis.

Confronting the challenges

Non-invasive prenatal diagnosis (NIPD) is a long sought-after goal in medical genetics. Currently, fetal genetic material must be collected through procedures such as amniocentesis and chorionic villus sampling to enable the definitive diagnosis of fetal genetic diseases. Unfortunately, these invasive procedures are associated with a risk of fetal loss. Although conventional teachings of reproductive biology state that such material should not be present, NIPD relies on the identification of traces of fetal genetic material in the blood of pregnant women. In 1969, Walknowska *et al.* [1,2] first reported the presence of fetal lymphocytes in maternal peripheral blood. Decades of research then followed in which intact fetal cells, present in the maternal circulation, were studied for use in NIPD [2]. However, the rarity of circulating fetal cells, typically several cells per milliliter of maternal whole blood, has prevented their robust detection [2]. At the conclusion of a ten-year multicenter study (1994–2003) funded by the United States National Institute of Child Health and Human Development, male fetal cells in whole maternal blood were detected in 41% of pregnancies with male fetuses, with a false-positive rate of 11% [3].

New opportunities for NIPD emerged in 1997 when cell-free fetal DNA was identified [4] and estimated to comprise some 10% of the total DNA in maternal plasma [5,6]. The development of several applications immediately followed and translated into clinical use [7,8], including in the determination of fetal gender for sex-linked disorders [4,9], fetal rhesus D blood group status in rhesus D negative women [7,8], and in the detection of paternally inherited mutations for autosomal dominant diseases [6,10]. However, the co-existence of a minor population of fetal DNA with the major background of maternal DNA in maternal plasma has posed challenges for extending the NIPD applications beyond those focusing on the detection of paternally inherited fetal alleles. Two such challenging areas included the achievement of NIPD for fetal chromosomal aneuploidies and monogenic diseases other than those caused by unique paternally inherited mutations.

In this review, we dissect the root causes of the challenges faced by NIPD researchers and review the attempts that have been taken to overcome them. In particular, we focus on the progress and efficacy in adopting the latest sophisticated analytical methods, namely digital polymerase chain reaction (PCR) and massively parallel sequencing, for non-invasive detection of fetal monogenic diseases and chromosomal aneuploidies. We conclude with comments on the present practical feasibility of these approaches and the possibility that prenatal diagnostic practices will be changed in the future.

Challenge 1: NIPD of monogenic diseases

Paternally inherited fetal alleles that are not shared by the maternal genome are distinguishable as fetal-specific in maternal plasma. Thus, detection of the presence or absence of paternally inherited mutations in maternal plasma can be readily applied to the NIPD of paternally inherited monogenic diseases. Reported examples included the NIPD of achondroplasia (a disorder which results in short stature), myotonic dystrophy and Huntington's disease [10]. However, it is much more challenging to achieve NIPD of maternally inherited or autosomal recessive monogenic diseases [6]. Maternally inherited fetal alleles are genotypically identical to the background maternal DNA; hence, fetal inheritance of a maternal mutation cannot be established by simply detecting its presence in maternal plasma. Similarly, for couples sharing identical mutations for an autosomal recessive condition, the fetal disease status cannot be assessed by mere detection of the mutation in maternal plasma.

Researchers have instead focused on the non-invasive prenatal exclusion of autosomal recessive diseases for

Corresponding author: Lo, Y.M.D. (loym@cuhk.edu.hk).

0168-9525/$ – see front matter doi:10.1016/j.tig.2009.05.004 Available online 18 June 2009

couples harboring non-identical mutations [11,12]. It is possible to determine either the absence of the paternal mutation in maternal plasma or to ascertain the presence of the paternal wild-type allele by detecting paternal-specific polymorphisms linked to the disease locus [11]. The aim of either strategy is to determine whether the fetus has inherited the paternal wild-type allele as this pattern excludes the chance of manifesting a disease which requires the inheritance of mutations from both parents. These strategies have been used for NIPD of β-thalassemia [11,13] and cystic fibrosis [12]. It is difficult to detect fetal-specific point mutations or single nucleotide polymorphisms (SNPs) with high analytical certainty when using standard tools such as real-time quantitative PCR (QPCR) [11,13,14]. Therefore, a challenge arises as just one nucleotide change in the targeted allele must be detected in the low fractional concentration of fetal DNA present in maternal plasma.

Challenge 2: NIPD of chromosomal aneuploidies

Down syndrome, where affected individuals typically have three instead of two copies of chromosome 21 (i.e. trisomy 21), is the most common aneuploidy, affecting ~1 in 700 births [15,16]. Other frequent aneuploidies include trisomy 18 (Edward syndrome), trisomy 13 (Patau syndrome) and monosomy X in females (Turner syndrome). Identification of trisomy 21 is the most common reason why women opt for prenatal diagnosis. Owing to the risk of fetal loss associated with invasive testing, methods involving ultrasonography and the analysis of maternal serum biochemical markers are currently used to risk-stratify those pregnancies which require confirmatory testing [17]. However, these screening tests detect the phenotypic features instead of the genetic pathology (i.e. root cause) of trisomy 21. The sensitivity and specificity profile of these tests is suboptimal, often requiring the combined use of multiple modalities, and they must be conducted within strict gestational age windows [17]. It would be ideal if direct detection of trisomy 21 could be achieved non-invasively.

Detection of chromosomal aneuploidy is a challenging puzzle in NIPD research. Fetal DNA in maternal plasma is cell-free. Thus, the dosage of chromosomes in the fetal genome cannot be determined as readily by methods such as fluorescence *in situ* hybridization. The high maternal DNA background also dilutes the genetic information one can obtain for the fetus through maternal plasma analysis. To overcome these issues, background maternal DNA interference can be minimized by the detection of molecular signatures that are present in maternal plasma but are contributed almost completely by the fetus. Circulating fetal DNA is derived predominantly from the placenta, whereas maternal DNA in plasma derives from maternal blood cells [18,19]. Genes that demonstrate differential DNA methylation [18] or expression profiles [20] between placental tissues and maternal blood cells have been developed as universal fetal nucleic acid markers for maternal plasma detection [10]. For example, serpin peptidase inhibitor, clade B (ovalbumin), member 5 (*SERPINB5*), also known as maspin [21,22], is hypomethylated in placental tissues but almost completely methylated in maternal blood cells [18]. Genotype analyses confirm that hypomethylated *SERPINB5* molecules in maternal plasma originate from the placenta or fetus.

SERPINB5 is located on chromosome 18. Tong *et al.* [23] achieved a chromosome 18 dosage comparison between trisomy 18 and euploid pregnancies by determining the ratio between polymorphic alleles of hypomethylated *SERPINB5* molecules in the maternal circulation. The rationale was based on the expectation that in a heterozygous trisomy 18 fetus, the ratio between the *SERPINB5* alleles would be 2:1 or 1:2 instead of 1:1 as in a heterozygous euploid fetus. Termed the epigenetic allelic ratio approach, it is the first strategy reported for the direct detection of chromosomal aneuploidies by cell-free DNA analysis from maternal plasma [23]. Extensive searches for chromosome 21 loci which demonstrate differential methylation between placental tissues and blood cells have been conducted [24–26] to extend the approach to the NIPD of trisomy 21. The low abundance of fetal DNA poses the main constraint on the practical feasibility of epigenetic based approaches [23].

Similarly, fetal-specific, placentally expressed mRNA molecules are detectable in maternal plasma. We developed an RNA–SNP allelic ratio test for the NIPD of trisomy 21 by determining the ratio between polymorphic alleles of placenta-specific 4 (*PLAC4*) mRNA, a transcript on chromosome 21, in maternal plasma [20]. *PLAC4* mRNA is more abundant than fetal DNA in maternal plasma and genotyping confirms that it is of fetal origin. Deviation of the *PLAC4* mRNA SNP allelic ratios were observed in plasma of trisomy 21 pregnancies compared with the expected 1:1 ratio in heterozygous euploid fetuses. 90% sensitivity and 96% specificity for the non-invasive detection of trisomy 21 were achieved by using the *PLAC4* RNA–SNP test alone; this is comparable to many of the currently used multi-modality screening tests.

Instead of simply targeting fetal-specific genetic signatures in maternal plasma, Dhallan *et al.* [27] attempted to reduce the maternal DNA background interference, thus resulting in a relative enrichment of fetal DNA. Formaldehyde was used as a cell stabilizing agent to minimize DNA release from maternal blood cells. Fetal aneuploidy was then detected by assessing statistically significant differences between polymorphic SNP ratios in chromosomes with and without involvement in the aneuploidy. However, controversies exist regarding the effectiveness of formaldehyde treatment because those findings could not be replicated consistently [28–32].

These approaches enable the direct detection of fetal chromosomal aneuploidies, but they are only applicable to fetuses with certain genotypes. For example, RNA–SNP tests are only informative for heterozygous fetuses; thus, a panel of coding SNPs is required to increase the population coverage of those tests.

Digitizing cell-free fetal DNA analysis

Most issues confounding circulating fetal DNA analysis are related to the interference caused by the high maternal DNA background. The approaches described rely on removing the influence of the background nucleic acids through the analysis of fetal-specific mutations or nucleic acid species

Box 1. Digital PCR

In digital PCR, template DNA is diluted to average concentrations of <1 molecule per well and analysed in hundreds to thousands of replicates [33]. Some PCR wells will be positive, whereas others will be negative for the targeted amplicon. Because most positive wells contain just one template molecule, counting the positive wells enables the absolute quantification of the original template DNA. Such quantification does not require the use of calibration standards or other gene targets for normalization; therefore, digital quantification is more accurate and precise than conventional QPCR [34,35]. In addition, by segregating template nucleic acid molecules into individual compartments, the amplification and detection of each template would not be affected by other templates with similar sequence context. As a result, the analytical power of digital PCR has been exploited in wide-ranging applications: the qualitative detection of trace molecular signatures, such as cancer mutations in heterogeneous biological samples [50]; quantitative imbalances between loci such as in loss of heterozygosity [51]; and copy number variation [52]. The performance of digital PCR is traditionally tedious and laborious, but disadvantages have been overcome by the introduction of nanofluidics devices for digital PCR analysis [35,52,53].

Box 2. Trisomy 21 detection

Fetal DNA co-exists with maternal DNA in maternal plasma. A trisomy 21 fetus, in comparison to a euploid fetus, adds extra amounts of chromosome 21 sequences into maternal plasma, in direct proportion to fetal DNA concentration. For example, a maternal plasma sample from a euploid pregnancy containing 100 genome-equivalents (GE)/ml of total DNA with 10 GE/ml DNA contributed by the fetus (i.e. 10% fetal DNA which is the typical median concentration for the first and second trimesters of pregnancy [35]) should contain a total of 200 copies (180 maternal copies + 20 fetal copies) of chromosome 21 sequences per milliliter of maternal plasma. For a trisomy 21 pregnancy, each fetal GE would contribute three copies of chromosome 21, resulting in a total of 210 copies/ml (180 maternal copies + 30 fetal copies) of chromosome 21 sequences in maternal plasma. At 10% fetal DNA concentration, the amount of chromosome-21-derived sequences in the maternal plasma of a trisomy 21 pregnancy would therefore be 1.05 times that of a euploid case. This degree of quantitative difference is difficult to discriminate confidently by QPCR [34].

in maternal plasma. However, such approaches do not provide general solutions. Methods need to be developed to extract fetal genetic information from circulating fetal DNA analysis in spite of co-existing maternal DNA molecules. Digital PCR presents one option [33,34] (Box 1).

By compartmentalizing individual template DNA molecules, digital PCR enables fetal and maternal DNA molecules in maternal plasma to be analysed separately without cross-interference. Specific detection of fetal alleles, for example point mutations or SNPs, in maternal plasma is therefore possible. Lun *et al.* [35] used a microfluidics digital PCR platform to detect fetal-derived Y-chromosomal DNA in maternal plasma. The fractional concentration of fetal DNA could be determined more accurately by this digital method than by conventional (analog) QPCR. The zinc finger protein homologs present on chromosomes Y (*ZFY*) and X (*ZFX*) were co-amplified using the same primer set to quantify the fetal and total DNA, respectively. The *ZFY* and *ZFX* amplicons differ by only two nucleotides; they are discriminated readily using duplex fluorogenic probes. In another study, Lun *et al.* [36] used duplex fluorogenic probes in a digital PCR assay to discriminate the wild-type hemoglobin, β (*HBB*) allele from a paternally inherited *HBB* point mutation, namely hemoglobin E, in maternal plasma. Previously, QPCR-based discrimination of fetal and maternal alleles that differed by just one or a few nucleotides was not specific or sensitive enough and required the use of more complex tools such as mass spectrometry [11,13,14]. Thus, the use of digital PCR could overcome the challenge of detecting fetal point mutations or SNPs in maternal plasma.

Quantitative power of molecule counting

In digital PCR (Box 1), direct counting of the wells with positive amplification of the target amplicon enables absolute quantification of the template DNA without the need for quantitative calibration standards. The analog and exponential nature of QPCR becomes a '1' and '0' signal in digital PCR [34]. Hence, digital counting platforms should enable more precise and accurate quantification [33,37]. For example, conventional QPCR can readily discriminate a difference in one or more threshold cycles (i.e. the number of cycles required for a reaction to reach a predetermined fluorescent threshold). As the amplicon concentration approximately doubles in each PCR cycle, the minimum quantitative difference that QPCR can easily discriminate is approximately a twofold change of template DNA concentration, which would be inadequate for detecting the increase in chromosome 21 DNA concentrations in maternal plasma of trisomy 21 pregnancies [38] (Box 2).

Using precise quantification methods, it should be possible to directly detect fetal aneuploidy by determining if the total (maternal + fetal) amount of the aneuploid chromosome (e.g. chromosome 21 for trisomy 21) is over-represented or under-represented compared with other chromosomes in maternal plasma. Digital PCR was performed for an amplicon located on chromosome 21 and another amplicon not on chromosome 21, that is, a reference chromosome [38,39]. The relative amounts of the two amplicons were compared in a strategy termed digital relative chromosome dosage (RCD) analysis [38]. Although over-representation of the chromosome 21 amplicon is expected in trisomy 21 fetuses, the degree of over-representation relies on the fetal DNA concentration (Box 2) and is smaller at low fetal DNA concentrations. Therefore, higher numbers of digital PCR analyses are required to ensure adequate statistical power to determine with confidence the presence or absence of chromosome 21 over-representation [38,39]. Our analysis of mixtures of placental and maternal blood cell DNA samples obtained from euploid and trisomy 21 pregnancies showed that NIPD of trisomy 21 could be accurately detected or excluded in 97% of cases by performing 7680 PCR analyses when the sample contained 25% fetal DNA [38]. Using mixtures of cell line DNA, Fan *et al.* [39] also demonstrated the use of digital PCR to detect chromosome 21 over-representation in trisomy 21.

Our study revealed several key parameters that affect digital RCD analysis of maternal plasma DNA [38]. Because the median concentration of circulating fetal DNA is usually <25% in first and second trimesters of

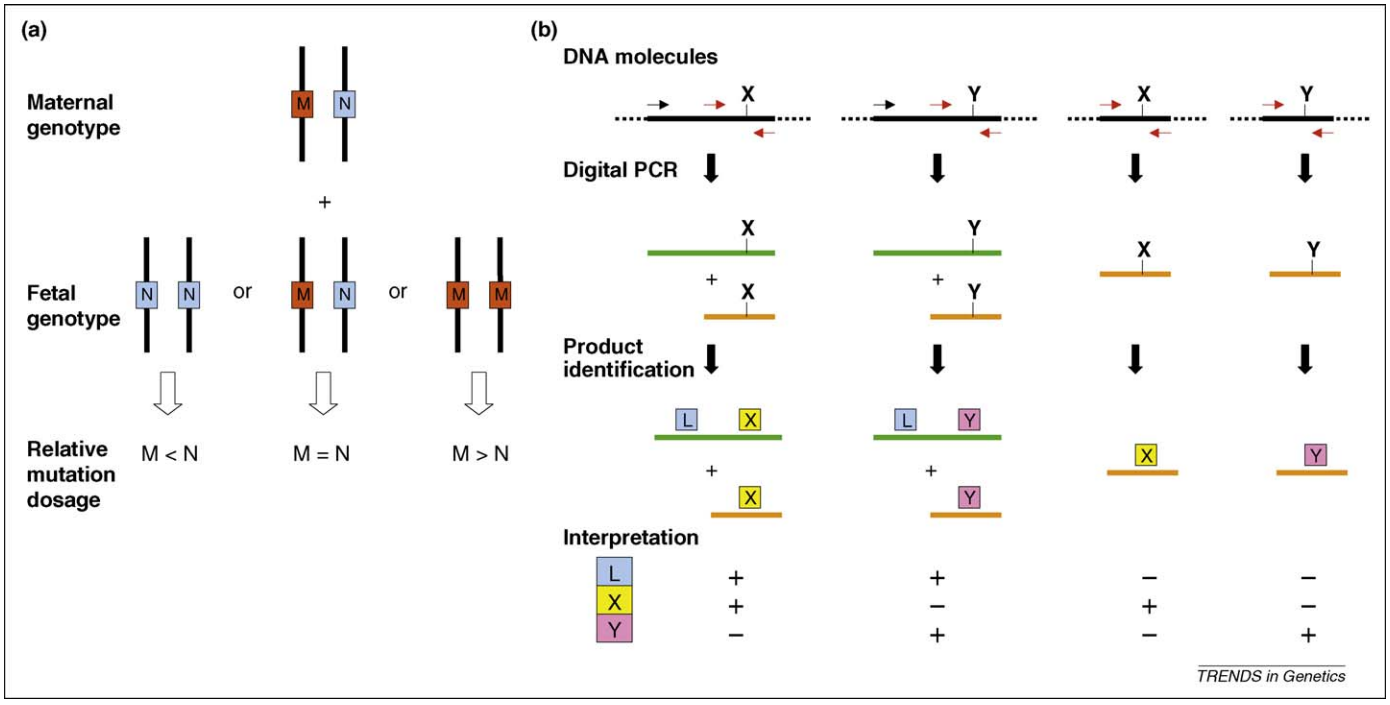

Figure 1. Schematic illustration of digital RMD and digital NASS. Digital RMD measures the relative amounts of the maternal mutant and wild-type alleles in maternal plasma to determine the inherited dosage of the mutant allele by the fetus. Digital NASS is a digital PCR-based method that enables the preferential analysis of short DNA molecules without physical size fractionation of DNA molecules. Lun *et al.* [36] applied the combined use of digital RMD and NASS to demonstrate the feasibility for the NIPD of fetal monogenic diseases. **(a)** When a pregnant woman and her fetus are both heterozygous for a gene mutation, the amounts of the mutant allele M (shown in orange) and wild-type allele N (shown in blue) are in allelic balance in maternal plasma. When the fetus is homozygous for the wild-type or mutant allele, there will be an under-representation or over-representation of the mutant allele, respectively. Digital RMD determines if the mutant and wild-type alleles in maternal plasma are in allelic balance or imbalance. **(b)** The scheme of the *ZFY–ZFX* digital NASS assay is shown. The assay can discriminate between *ZFX*, denoted by X, and *ZFY*, denoted by Y, DNA molecules. In addition, the assay can distinguish if the *ZFX* and *ZFY* DNA molecules are long or short. Digital PCR is performed using two forward primers (black and red arrows pointing right) and one reverse primer (red arrows pointing left), or vice versa, that are oriented to produce a short amplicon overlapping with the long amplicon. When a single DNA molecule at least as long as that specified by the long amplicon is captured in the reaction well, both the long (green lines) and short (orange lines) PCR products are generated. When a single DNA molecule shorter than the span of the long amplicon is captured, only the short amplicon is generated. The presence of the long and/or short amplicons can be detected by strategically located hybridization probes or extension primers. An extension primer, L (blue box), is designed to detect the presence of the long amplicon. A separate extension primer is located within the short amplicon, and the extension products are used to discriminate the *ZFX* and *ZFY* alleles (shown as boxed X [yellow] and boxed Y [pink], respectively). The identities of the DNA molecules can be determined by counting the products present within each well. +, present; −, absent. Figure adapted, with permission, from Ref. [36] (Copyright, 2008; National Academy of Sciences, U.S.A.).

pregnancy [35], either fetal DNA enrichment or additional PCR analyses are required. Indeed, if the fetal DNA concentration is halved, four times as many digital PCR analyses are needed [38]. Because maternal plasma DNA concentrations are typically thousands of copies/mL, tens of milliliters of maternal blood are needed to perform tens of thousands of digital PCRs. Alternatively, given that plasma DNA exists as small fragments, instead of targeting one chromosome 21 amplicon, digital detection of multiple multiplexed amplicons on chromosome 21 would effectively increase the number of digital PCR datapoints within the same fixed maternal plasma volume [38]. Further studies are required to evaluate the efficacy of such options.

From relative chromosome dosage to relative mutation dosage

Principles similar to digital RCD have been developed for NIPD of monogenic diseases. Instead of targeting only paternal mutations, it is possible to compare the relative amounts of the maternal mutant and wild-type alleles in maternal plasma to determine the inherited dosage of the mutant allele. Termed the relative mutation dosage (RMD) approach [36,40], this application is most clinically relevant for pregnant women who are heterozygous for a given mutation (Figure 1a). If the fetus has not inherited the mutation, under-representation of the mutant allele is expected. If the fetus is homozygous for the mutation (i.e. a second mutant copy was contributed by the father), over-representation of the mutant allele is expected. Lastly, if the fetus is heterozygous for the maternal mutation, the mutant and wild-type alleles should be in allelic balance. Lun *et al.* [36] used digital RMD maternal plasma analysis to determine the fetal inheritance of hemoglobin E and β-thalassemia mutations in mothers who are carriers of either mutation. The principles of digital RMD are feasible but, as with digital RCD, large numbers of digital PCR analyses are required for samples containing low fetal DNA concentrations. To render digital RMD more practical, it was combined with a fetal DNA enrichment strategy [36]. Fetal DNAs are shorter than maternal DNAs [41] and, therefore, size fractionation of short DNA molecules can enrich fetal DNA. Instead of using physical methods of size fractionation, such as gel electrophoresis [13], Lun *et al.* [36] used a digital method, termed digital nucleic acid size selection (NASS), to derive information from short DNA molecules. Digital NASS uses a duplex digital PCR assay targeting overlapping amplicons of different sizes (Figure 1b). During NASS analysis, only wells showing the presence of short DNA molecules are counted for RMD

assessment. In maternal plasma samples, the combined use of digital NASS and RMD enables the fetal genotype to be discernible in cases in which RMD alone would be insufficient [36]. With these new developments, NIPD of paternally or maternally contributed or autosomal recessive monogenic diseases can be achieved.

Molecule counting by massively parallel maternal plasma DNA sequencing

The low fetal DNA fractional concentration and the low absolute concentration of template DNA in maternal plasma requires either fetal DNA enrichment or large numbers of counted DNA molecules to bring NIPD by molecule counting close to clinical use. As plasma DNA is fragmented, instead of targeting specific loci in the genome, a locus-independent method could be used. One copy of chromosome 21 therefore would be sampled and counted many times in a locus-independent method, instead of just once, as for example in locus-specific assays (Figure 2). By increasing the number of measurements per sample, higher analytical precision can be achieved without the need to increase the volume of input maternal plasma. The recent availability of massively parallel sequencing platforms [42] have been adopted as a tool for maternal plasma DNA analysis for the NIPD of trisomy 21 and potentially other chromosomal aneuploidies [43,44].

The rationale is to use massively parallel sequencing to count DNA molecules in maternal plasma. When a woman is pregnant with a trisomy 21 fetus, an over-representation of the fractional concentration of chromosome 21 sequences in her plasma is expected (Box 2). Therefore, if a random representative portion of DNA fragments from a maternal plasma sample is sequenced, the frequency distribution of the chromosomal origin of the sequenced DNA fragments should reflect the genomic representation of the original maternal plasma sample. Assuming that the genomic representations of maternal and fetal DNA in maternal plasma are not grossly skewed or uneven across the chromosomes, an increased proportion of chromosome 21 sequences should be present in relation to the total sequenced reads for DNA in maternal plasma obtained from a trisomy 21 pregnancy.

Both Chiu *et al.* [43] and Fan *et al.* [44] demonstrated the use of massively parallel plasma DNA sequencing for NIPD of trisomy 21 on Illumina's Genome Analyzer platform [45]. A short region on one end of each plasma DNA molecule is sequenced and aligned computationally to the reference human genome to determine the chromosomal origin of each DNA fragment. The proportion of sequenced reads from chromosome 21 were compared between trisomy 21 and euploid pregnancies. Fan *et al.* [44] used 1.3 mL to 3.2 mL plasma from nine trisomy 21 and six euploid pregnancies and obtained ~5 million unique reads per sample, with up to one nucleotide mismatch. The sequence density per 50-kb window for chromosome 21 was normalized by the median value obtained from the euploid cases. The normalized sequence tag densities for the trisomy 21 cases were >99% confidence interval bound for the euploid cases.

Chiu *et al.* [43] used DNA from 5 mL to 10 mL plasma from 14 trisomy 21 and 14 euploid pregnancies and obtained a mean of ~2 million unique reads per sample, without mismatches to the reference human genome. The number of reads originating from chromosome 21 was expressed as a proportion of all sequenced reads, and z-scores, representing the number of standard deviations away from the mean proportion of chromosome 21 reads in a reference set of euploid cases, were determined for each case. A z-score $>\pm 3$ indicated a 99% chance of a statistically significant difference in the assessed parameter for the test case compared with the reference group. Thus, a high z-score was expected for trisomy 21 cases. The mas-

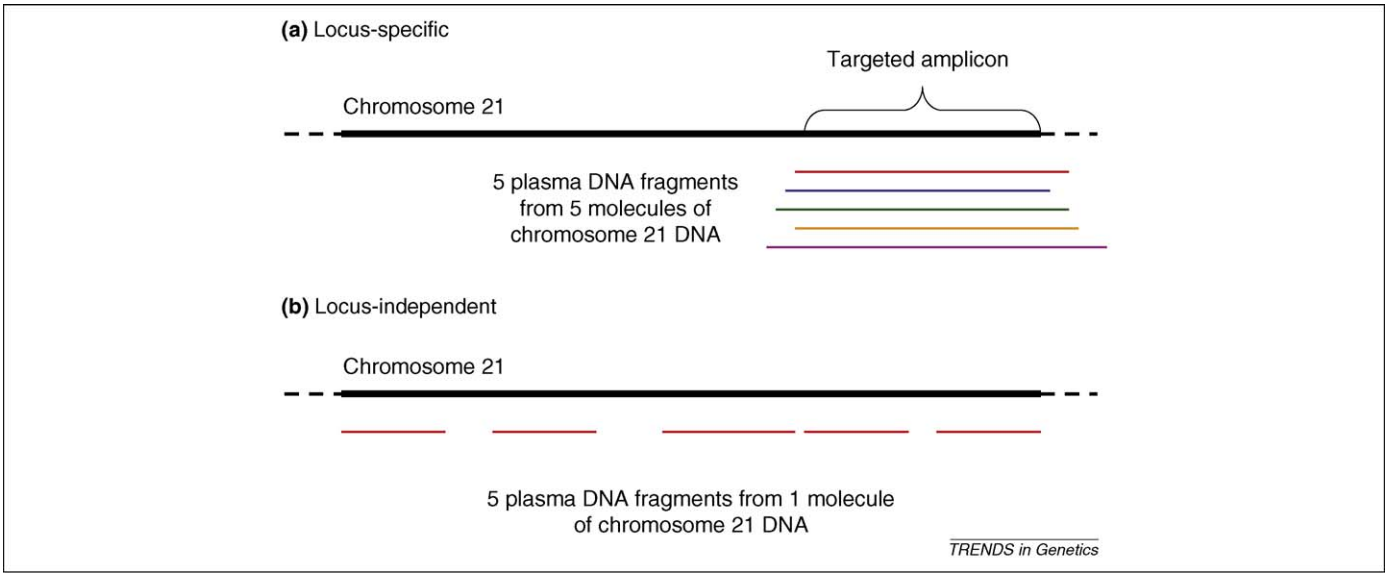

Figure 2. Schematic comparison between locus-specific and locus-independent methods for DNA quantification. DNA molecules exist as short fragments in maternal plasma [41]. Hence, instead of comparing the relative amounts between specific loci as with conventional DNA quantification methods [5], the amount of quantitative information that one could derive with the same amount of plasma DNA input greatly increases with the use of locus-independent quantification methods that treat each DNA fragment as an individual target. **(a)** When using locus-specific assays, five copies of chromosome 21 with the targeted amplicon region intact (depicted by the different colored DNA molecules) would need to be physically present to generate a count of five. **(b)** In the locus-independent method, five fragmented portions originating from a single chromosome 21 (depicted by the DNA fragments of the same color) could potentially contribute to a count of five.

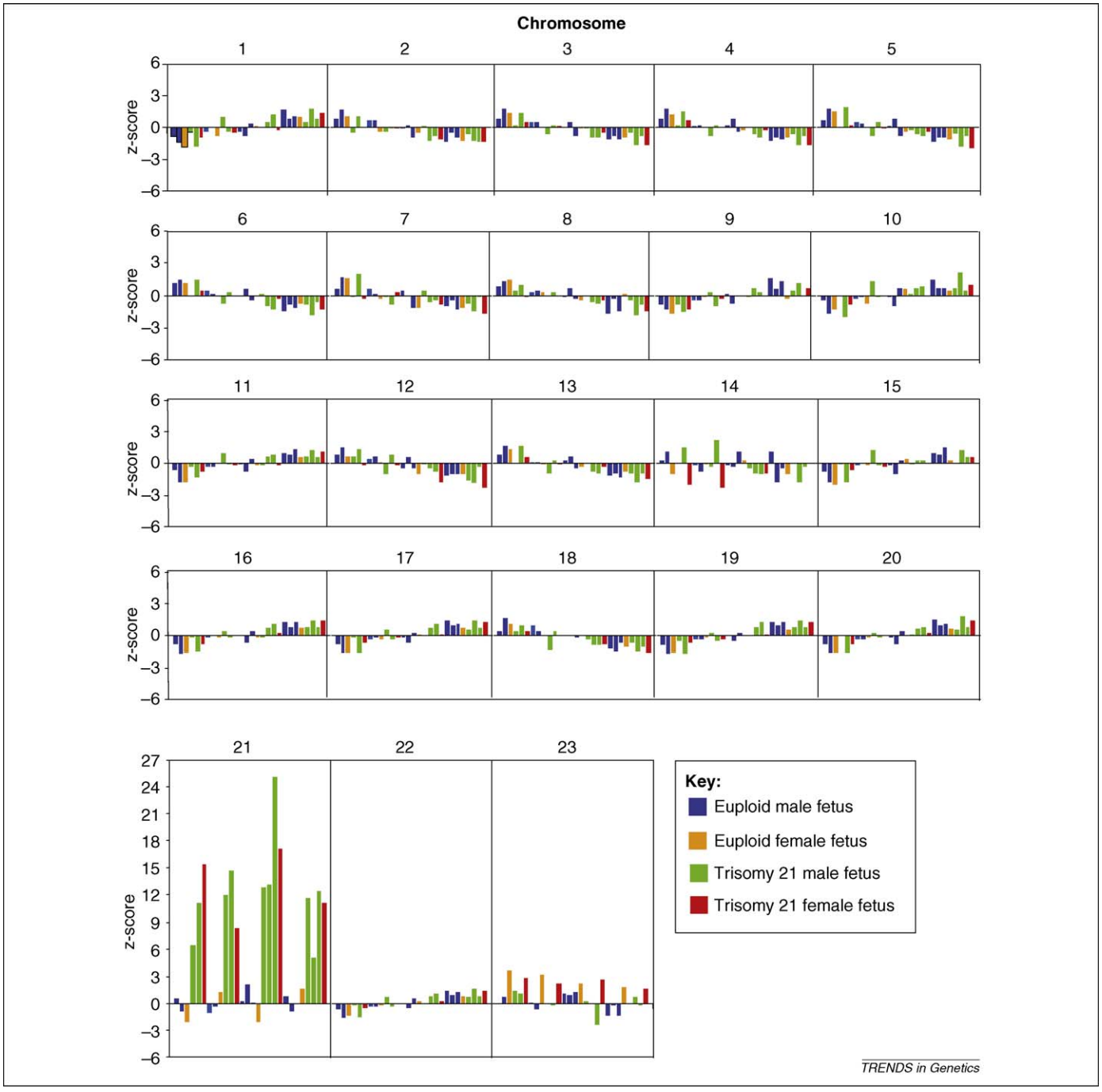

Figure 3. Detection of fetal trisomy 21 by massively parallel sequencing. In this approach, a random representative portion of DNA molecules in maternal plasma is sequenced. The chromosomal origin of each sequenced read is identified by bioinformatics analysis. The mean and standard deviation of the proportion of reads from each chromosome of a reference sample set comprising pregnancies with euploid male fetuses are determined. Z-scores, representing the number of standard deviations from the mean of the reference sample set, of the percentage chromosomal representation for each maternal plasma sample are calculated. Z-scores beyond $\pm$ 3 suggest a >99% chance of the presence of chromosomal over- or under-representation compared with the reference group. Here, plots of z-scores for each chromosome for maternal plasma samples from 14 trisomy 21 and 14 euploid pregnancies are shown. Each of the 28 bars shown for each chromosome corresponded to the z-scores for one of the 28 maternal plasma samples. Samples 1 to 28 are shown consecutively from left to right. Figure adapted, with permission, from Ref. [43] (Copyright, 2008; National Academy of Sciences, U.S.A.).

sively parallel sequencing approach was reliable and robust: in all cases, z-scores $<\pm$ 3 were obtained for all chromosomes except 21 and X (Figure 3). Z-scores of chromosome 21 were beyond +5 for all 14 trisomy 21 cases but within $\pm$ 3 for all euploid cases. Because pregnancies with male fetuses were used as the reference sample set, z-scores for the X-chromosome were increased in all pregnancies with female fetuses.

Both studies demonstrated that massively parallel sequencing can randomly count and identify DNA fragments in maternal plasma in a locus-independent manner (Figure 2) to detect small quantitative perturbations in genomic distribution of plasma DNA [43,44]. The large number of measurements done per sample enables a highly precise estimation of the proportion of chromosome 21 sequences; hence, its over-representation in trisomy 21

can be robustly detected. The robustness of the approach further suggests that the genomic distributions of maternal and fetal DNA molecules in maternal plasma are unlikely to be grossly skewed. In spite of the low abundance of fetal DNA in maternal plasma [5,35], fetal trisomy 21 was detected accurately by sequencing. We suspect that because fetal DNA molecules are shorter than the maternal DNA background and present more 'ends' for short-read sequencing, there might be a degree of enrichment in reads originating from the fetal genome. However, both reports [43,44] observed that the imprecision for measuring the genomic representation of chromosomes is variable and depends on the GC content of each chromosome. For example, the genomic representations of chromosomes 18 and 13, which are relevant for trisomy 18 and trisomy 13, respectively, cannot be measured as precisely as chromosome 21 [43].

Fan *et al.* [44] collected maternal plasma samples (except one collected at 35 weeks gestation) 15 to 30 min after amniocentesis, a second-trimester procedure, or chorionic villus sampling, a first-trimester procedure. Within this time frame, maternal plasma fetal DNA concentrations have been reported to be elevated as a result of the invasive procedures [46]. In addition, circulating fetal DNA concentrations increase with pregnancy progression [5,10]. However, the median gestational ages for the trisomy 21 (18 weeks) and the euploid (12 weeks) groups in the study by Fan *et al.* [44] were within different trimesters. This difference could enhance the chance of detecting increases in fetal DNA contribution to the sequenced reads in the trisomy 21 group [44]. By contrast, clinical samples analysed by Chiu *et al.* [43] were predominantly collected before invasive procedures with comparable median gestational ages for the euploid and trisomy 21 groups. In a minority of cases, maternal blood was collected days after invasive procedure at a time which fetal DNA is known to return to baseline concentrations [43]. The reliability of the massively parallel sequencing approach for NIPD of trisomy 21 is promising, but confirmations from carefully controlled, larger scale studies are needed.

Concluding remarks: counting down towards routine NIPD

The amount of information that can be obtained from each maternal plasma sample through the use of massively parallel sequencing is phenomenal. Depending on the performance and cost profiles of the newly developed tests, NIPD of fetal chromosomal aneuploidies could be implemented as a secondary screening or diagnostic test for high-risk pregnancies identified by the current screening programs, or ultimately for all pregnancies as a first line test. With the availability of tools to selectively enrich portions of the genome for resequencing [42], massively parallel sequencing could be applicable to the NIPD of monogenic diseases. Alternatively, the depth of sequencing could be increased, along with the costs, until locus-specific mutations are captured by sequencing with high certainty. Although controversial, one day the approach might enable non-invasive deciphering of the complete fetal genome during the antenatal period. Massively parallel sequencing is potentially useful for molecular diagnostics involving plasma DNA analysis for other clinical settings [47] including the study of tumor-derived DNA [48] and infectious diseases [49].

Currently, a major limitation to the implementation of sequencing-based plasma diagnostics is its high cost. Compared with QPCR and mass spectrometry, the equipment is expensive, the reagent costs are high and the present throughput is low. For example, the published protocols [43,44] enable the processing of eight plasma samples per 3–5 day run. Barcoding of DNA samples could increase the throughput and decrease the costs, thereby enabling tens to even hundreds of samples to be analysed per run. Nonetheless, the costs for sequencing are expected to continue to decrease sharply until they are no longer prohibitory to its routine implementation. Until such a time, combined use of sequencing with the more economical approaches, such as RNA–SNP for trisomy 21, is a viable option for bringing forth the clinical implementation of NIPD.

This is an exciting time for NIPD research. Single molecule counting techniques enable analysis and quantification of nucleic acids at unprecedented sensitivity, accuracy and precision. These powerful new tools have the potential of bringing plasma nucleic-acid-based diagnostics another step closer to widespread clinical use.

Disclosure statement

The authors have filed patent applications on the detection of fetal nucleic acids in maternal plasma for non-invasive prenatal diagnosis. Part of this patent portfolio has been licensed to Sequenom. C.R.C. is the Chief Scientific Officer of, and holds equities in, Sequenom. Y.M.D.L. is a consultant to, and holds equities in, Sequenom.

Acknowledgements

Y.M.D.L. and R.W.K.C. were supported by the University Grants Committee of the Government of the Hong Kong Special Administration Region, China, under the Areas of Excellence Scheme (AoE/M-04/06) and a sponsored research agreement with Sequenom. Y.M.D.L. was supported by an Endowed Professorship from the Li Ka Shing Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1 Walknowska, J. *et al.* (1969) Practical and theoretical implications of fetal-maternal lymphocyte transfer. *Lancet* 1, 1119–1122
2 Bianchi, D.W. (2004) Fetomaternal cell traffic, pregnancy-associated progenitor cells, and autoimmune disease. *Best Pract. Res. Clin. Obstet. Gynaecol.* 18, 959–975
3 Bianchi, D.W. *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
4 Lo, Y.M.D. *et al.* (1997) Presence of fetal DNA in maternal plasma and serum. *Lancet* 350, 485–487
5 Lo, Y.M.D. *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
6 Lo, Y.M.D. and Chiu, R.W.K. (2007) Prenatal diagnosis: progress through plasma nucleic acids. *Nat. Rev. Genet.* 8, 71–77
7 Daniels, G. *et al.* (2009) Noninvasive prenatal diagnosis of fetal blood group phenotypes: current practice and future prospects. *Prenat. Diagn.* 29, 101–107
8 Van der Schoot, C.E. *et al.* (2006) Non-invasive antenatal RHD typing. *Transfus. Clin. Biol.* 13, 53–57
9 Costa, J.M. *et al.* (2002) New strategy for prenatal diagnosis of X-linked disorders. *N. Engl. J. Med.* 346, 1502

10 Wright, C.F. and Burton, H. (2009) The use of cell-free fetal nucleic acids in maternal blood for non-invasive prenatal diagnosis. *Hum. Reprod. Update* 15, 139–151
11 Ding, C. *et al.* (2004) MS analysis of single-nucleotide differences in circulating nucleic acids: Application to noninvasive prenatal diagnosis. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10762–10767
12 Nasis, O. *et al.* (2004) Improvement in sensitivity of allele-specific PCR facilitates reliable noninvasive prenatal detection of cystic fibrosis. *Clin. Chem.* 50, 694–701
13 Li, Y. *et al.* (2005) Detection of paternally inherited fetal point mutations for β-thalassemia using size-fractionated cell-free DNA in maternal plasma. *JAMA* 293, 843–849
14 Tsang, J.C.H. *et al.* (2007) Mass spectrometry-based detection of hemoglobin E mutation by allele-specific base extension reaction. *Clin. Chem.* 53, 2205–2209
15 Sherman, S.L. *et al.* (2007) Epidemiology of Down syndrome. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 221–227
16 Wiseman, F.K. *et al.* (2009) Down syndrome–recent progress and future prospects. *Hum. Mol. Genet.* 18, R75–R83
17 Malone, F.D. *et al.* (2005) First-trimester or second-trimester screening, or both, for Down's syndrome. *N. Engl. J. Med.* 353, 2001–2011
18 Chim, S.S.C. *et al.* (2005) Detection of the placental epigenetic signature of the maspin gene in maternal plasma. *Proc. Natl. Acad. Sci. U. S. A.* 102, 14753–14758
19 Alberry, M. *et al.* (2007) Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. *Prenat. Diagn.* 27, 415–418
20 Lo, Y.M.D. *et al.* (2007) Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat. Med.* 13, 218–223
21 Dokras, A. *et al.* (2002) The tumour suppressor gene maspin is differentially regulated in cytotrophoblasts during human placental development. *Placenta* 23, 274–280
22 Futscher, B.W. *et al.* (2002) Role for DNA methylation in the control of cell type specific maspin expression. *Nat. Genet.* 31, 175–179
23 Tong, Y.K. *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
24 Old, R.W. *et al.* (2007) Candidate epigenetic biomarkers for non-invasive prenatal diagnosis of Down syndrome. *Reprod. Biomed. Online* 15, 227–235
25 Chim, S.S.C. *et al.* (2008) Systematic search for placental DNA-methylation markers on chromosome 21: toward a maternal plasma-based epigenetic test for fetal trisomy 21. *Clin. Chem.* 54, 500–511
26 Papageorgiou, E.A. *et al.* (2009) Sites of differential DNA methylation between placenta and peripheral blood: molecular markers for noninvasive prenatal diagnosis of aneuploidies. *Am. J. Pathol.* 174, 1609–1618
27 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
28 Benachi, A. *et al.* (2005) Impact of formaldehyde on the *in vitro* proportion of fetal DNA in maternal plasma and serum. *Clin. Chem.* 51, 242–244
29 Chinnapapagari, S.K. *et al.* (2005) Treatment of maternal blood samples with formaldehyde does not alter the proportion of circulatory fetal nucleic acids (DNA and mRNA) in maternal plasma. *Clin. Chem.* 51, 652–655
30 Chung, G.T.Y. *et al.* (2005) Lack of dramatic enrichment of fetal DNA in maternal plasma by formaldehyde treatment. *Clin. Chem.* 51, 655–658
31 Hulten, M.A. and Old, R.W. (2007) Non-invasive prenatal diagnosis of Down's syndrome. *Lancet* 369, 1997
32 Lo, Y.M.D. and Chiu, R.W.K. (2007) Non-invasive prenatal diagnosis of Down's syndrome. *Lancet* 369, 1997
33 Vogelstein, B. and Kinzler, K.W. (1999) Digital PCR. *Proc. Natl. Acad. Sci. U. S. A* 96, 9236–9241
34 Pohl, G. and Shih Ie, M. (2004) Principle and applications of digital PCR. *Expert Rev. Mol. Diagn.* 4, 41–47
35 Lun, F.M.F. *et al.* (2008) Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin. Chem.* 54, 1664–1672
36 Lun, F.M.F. *et al.* (2008) Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19920–19925
37 Wang, T.L. *et al.* (2002) Digital karyotyping. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16156–16161
38 Lo, Y.M.D. *et al.* (2007) Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc. Natl. Acad. Sci. U. S. A.* 104, 13116–13121
39 Fan, H.C. and Quake, S.R. (2007) Detection of aneuploidy with digital polymerase chain reaction. *Anal. Chem.* 79, 7576–7579
40 Zimmermann, B.G. *et al.* (2008) Digital PCR: a powerful new tool for noninvasive prenatal diagnosis? *Prenat. Diagn.* 28, 1087–1093
41 Chan, K.C.A. *et al.* (2004) Size distributions of maternal and fetal DNA in maternal plasma. *Clin. Chem.* 50, 88–92
42 Voelkerding, K.V. *et al.* (2009) Next-generation sequencing: from basic research to diagnostics. *Clin. Chem.* 55, 641–658
43 Chiu, R.W.K. *et al.* (2008) Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20458–20463
44 Fan, H.C. *et al.* (2008) Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc. Natl. Acad. Sci. U. S. A.* 105, 16266–16271
45 Bentley, D.R. *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59
46 Samura, O. *et al.* (2003) Cell-free fetal DNA in maternal circulation after amniocentesis. *Clin. Chem.* 49, 1193–1195
47 Lo, Y.M.D. and Chiu, R.W.K. (2009) Next-generation sequencing of plasma/serum DNA: an emerging research and molecular diagnostic tool. *Clin. Chem.* 55, 607–608
48 Korshunova, Y. *et al.* (2008) Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. *Genome Res.* 18, 19–29
49 Beck, J. *et al.* (2009) Profile of the circulating DNA in apparently healthy individuals. *Clin. Chem.* 55, 730–738
50 Diehl, F. *et al.* (2005) Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16368–16373
51 Zhou, W. *et al.* (2001) Counting alleles reveals a connection between chromosome 18q loss and vascular invasion. *Nat. Biotechnol.* 19, 78–81
52 Qin, J. *et al.* (2008) Studying copy number variations using a nanofluidic platform. *Nucleic Acids Res.* 36, e116
53 Morrison, T. *et al.* (2006) Nanoliter high throughput quantitative PCR. *Nucleic Acids Res.* 34, e123