Noninvasive prenatal detection of genetic defects

Oever, Jessica Maria Elisabeth van den

Chapter 7

Discussion

Discussion

Fragmented cell-free fetal DNA (cffDNA) present in maternal plasma is a potential source for prenatal genetic analysis of the fetus early in gestation. Although the amount of fetal fragments in maternal plasma is relatively low as compared to the excess of maternal DNA, results from the study of Lo *et al.* showed that the entire fetal genome is present in maternal plasma (Lo *et al.*, 2010). In theory, this could mean that every genetic defect in the fetus can be detected in maternal plasma by noninvasive prenatal diagnosis (NIPD). Since the whole fetal genome is present, a genome wide approach for NIPD could be used. Shotgun Whole genome sequencing (WGS) has already been applied for NIPT. However, for NIPD, shotgun WGS currently has its limitations. Many millions of sequencing reads can be obtained genome wide. Yet, shotgun sequencing occurs randomly and is not mutation specific. Hence, the coverage of a mutation or region of interest may be too low for reliable NIPD. A universal approach, such as WGS, for NIPD may therefore not be achievable or the preferred method of choice yet. Novel applications as well as other approaches could be considered and explored. There is a growing demand for less invasive alternatives in prenatal diagnostics. Therefore, it is to be expected that in the near future the number of available tests will increase dramatically. In this thesis, we have described several applications that make use of cfDNA from maternal plasma in NIPD and NIPT, such as fetal trisomy screening and targeted detection of paternally inherited mutations.

Fetal markers

In both NIPD and NIPT determination of the fetal fraction in maternal plasma can be important. Male specific markers (e.g. *DYS14* and *SRY*) are used most often for this purpose. Despite the high sensitivity and specificity obtained in diagnostic tests with these markers, a positive result can only be obtained in case of a male fetus. Preferably a universal and sex independent marker such as fetal methylated *RASSF1A* (*mRASSF1A*) should be used to confirm the presence of cffDNA in maternal plasma. This universal marker is differentially expressed between mother and fetus. In **chapter 2** we have described a novel approach for the detection of *mRASSF1A* in maternal plasma. We have used a combination of bisulfite conversion and pyrophosphorolysis-activated polymerization (PAP) for the detection of *mRASSF1A*. Bisulfite conversion is necessary to reveal the differences in methylation of cytosines between maternal and fetal sequences. PAP is an extremely sensitive and specific technique which can be used to detect specific sequences within a high background, in this case fetal specific sequences within the maternal background. Previously, the use of PAP in NIPD has proven to be of value in fetal sexing (Boon *et al.*, 2007). PAP has also been described for the detection of residual tumour cells in blood, serum or plasma (Maat *et al.*, 2008). *RASSF1A* is a tumour suppressor gene and *RASSF1A* promoter hypermethylation is associated with loss of expression in tumour cells (Honorio *et al.*, 2003). The fact that *RASSF1A* expression is also tumour related is a very important consideration. Unknown underlying malignancies or circulating tumour cells in the maternal circulation may result in false positive results, although the prevalence of cancer in the reproductive age group is considered to be small. The laboratory should therefore always be informed when women opting for NIPD have a history of cancer. Other alternative tests could then be performed to confirm the presence of cffDNA in plasma, such as the determination of paternally inherited SNPs (Alizadeh *et al.*, 2002; Page-Christiaens *et al.,* 2006). This assay is a good, although labour intensive, alternative for the PAP assay. Also, one should keep

in mind that SNPs may not always be informative. Parents could have all tested SNPs in common, or the fetus did not inherit an informative paternal SNP.

The *mRASSF1A*-PAP we currently perform for fetal sexing is merely a qualitative test, rather than quantitative. Quantitative detection of a sex independent fetal marker would be extremely beneficial in both NIPD and NIPT to determine the fetal fraction in a sample. A digitalized quantitative PAP assay has previously been described, though these assays were performed without prior bisulfite conversion of the DNA (Madic *et al.*, 2012; Song *et al.*, 2014; Bidard *et al.*, 2014). It has to be elucidated whether this combination of bisulfite conversion and subsequent digitalized PAP for the detection of *mRASSF1A* in maternal plasma would work. It would be very advantageous to combine such a digitalized PAP assay with NIPD of monogenic disorders. Not only the fetal paternally inherited mutation could then be detected, also the amount of cffDNA in the sample could be determined to define fetal fraction using the same sample. This in contrast to NIPT, where the amount of fetal DNA is determined either prior or after sample prep. At present, it is not yet clear which one of these determinations is the best predictor for the actual fetal fraction of a sample, since it is also not yet clear how these results will best represent the fetal fraction in the actual NGS data.

In the last few years, the search for other differentially methylated regions (DMRs) has continued (Papageorgiou *et al.*, 2009; Ioannides *et al.*, 2014; Xiang *et al.*, 2014). Most of this research is performed with methylated DNA immunoprecipitation sequencing (MeDIP-Seq) together with comparative arrays to provide global methylation landscapes. The differences in methylation between mother and fetus are based on relative differences in expression or methylation level. Some candidate DMRs show a high level of heterogeneity between individual fetuses. Yet, good candidate DMRs for use in NIPT or NIPD require a high level of homogeneity between individual fetuses and should be able to exclude maternal DNA easily (Xiang *et al.*, 2014). Even though MeDIP-Seq will give a good indication of which regions are differentially methylated, for use in a PAP assay, it has to be confirmed at nucleotide level whether these regions are indeed good candidate DMRs. The focus for the discovery of novel DMRs has been on chromosome 21 and 18. These DMRs have mostly been used for fetal trisomy screening (Tong *et al.*, 2006; Chim *et al.*, 2008; Tong *et al.*, 2010a; Tong *et al.*, 2010b). Furthermore, recent studies have shown the determination of the fetal methylome and transcriptome (Lun *et al.*, 2013; Tsui *et al.*, 2014). In the near future, perhaps more candidate DMRs will be discovered that can be useful for NIPD or NIPT.

NIPT

For NIPT, we have proposed Helicos Single Molecule Sequencing (SMS) as an alternative non-PCR-based sequencing platform in **chapter 3 and 4**. In contrast to PCR-based platforms, Helicos SMS data show no GC bias. In this study, SMS was successful for both T21 and T18 detection. For T13 however, it appeared less optimal. In several NIPT studies performed on PCR-based platforms, the detection rate for NIPT of fetal trisomy 13 was reported to be lower as compared to T21 and T18 detection (Gil *et al.*, 2014). In several other studies, the lower detection rate has been explained by the difference in GC content between these chromosomes, mainly because chromosome 13 has the lowest average GC content of these three chromosomes (Chen *et al.*, 2011; Bianchi *et al.*, 2012; Lau *et al.*, 2012b; Palomaki *et al.*, 2012; Norton *et al.*, 2012; Sparks *et al.*, 2012a). In chapter 3 we have shown that the mean number of reads per 50 kb bin increases when GC content increases on a PCR based sequencing platform. For Helicos the mean read count was not influenced by GC content, thus Helicos data display no GC bias. Therefore, the difference in GC content could not explain the lower detection rate of

T13 with SMS. Kalousek *et al.* showed that all placentas from T13 and T18 fetuses examined in their study were mosaic (Kalousek *et al.*, 1996). In case of such confined placental mosaicism (CPM), the presence of large quantities of cells with different karyotypes may influence results of fetal aneuploidy detection. Even though CPM occurs frequently in T13 and T18 placenta's, in our preliminary study using SMS possible CPM could also not explain lower performance observed for T13, since full karyotyping of the samples showed CPM was not present. It should however be noted, that the number of samples tested for these preliminary studies using SMS has been low. Extending this data set would be very useful to elucidate the reason for the lower detection rate of T13 screening and to investigate the role of GC content and/or bias for this particular chromosome. Unfortunately, since 2012 Helicos services were no longer available because of bankruptcy of the company and these tests could therefore not be performed for this thesis. A revival of the company might open possibilities to run additional samples to extend this data for fetal trisomy screening by use of SMS.

We have compared Helicos SMS with the Illumina GAII platform, which was mostly used for noninvasive fetal trisomy screening at that time. Since then, many improvements in the Illumina sequencing technology have been extremely beneficial for NIPT purposes and at present, Illumina sequencers (e.g. HiSeq) are most frequently used for NIPT. Newer updates with features such as a rapid run modus and decreased turn-around time are extremely beneficial for NIPT. With the introduction of novel sequencers such as HiSeq X Ten, which combines Nano technology to obtain sufficient data for a few genomes per single run, the costs for running NIPT samples may even further be reduced (Buermans *et al.*, 2014; Hayden, 2014).

Calculation methods and downstream analysis pipelines for NIPT have been improved in the last few years. Studies that use ratio calculations for fetal trisomy detection as described in chapter 3 and 4 have not been reported since. Currently the Z-score, NCV (**Appendix 3**) and student's t-test-based methods are mostly used, either as stand-alone methods or combined (Chiu *et al.*, 2008; Sehnert *et al.*, 2011; Jiang *et al.*, 2012a; Lo *et al.*, 2014; Straver *et al.*, 2014; Bayindir *et al.*, 2015). Thousands of NIPT samples have been processed in the past few years by many different groups and all these groups have been putting a lot of effort into optimizing NIPT for use in their laboratory setting, most frequently using the Illumina platform for sequencing.

Other single molecule sequencing platforms beside Helicos SMS have been developed in the past years, although not every platform may be as suitable for NIPT purposes. The PacBio sequencer by Pacific Biosciences is a single molecule real-time (SMRT) DNA sequencing system that records light pulses emitted as a by-product of nucleotide incorporation. One of the most important features of SMRT sequencing is fast run time (within hrs) and the extremely long read length that can be obtained. Compared to the 8-9 day running time of the Helicos flow cells, the PacBio sequencer is faster and scalable in runtime, which is favourable for NIPT turn-around time. Depending upon PacBio starting library, over half of the data are in reads > 14,000 base pairs long, starting from around 3000 bp with the longest reads over 40,000 bp. Such long read lengths can indeed be beneficial for several applications, nevertheless cfDNA in maternal plasma is on average a hundredfold shorter than the reads that can be obtained with this platform. Therefore, PacBio is not a suitable platform for NIPT. Also Oxford Nanopores, a platform that uses nanopores as biosensors, has similar advantages (e.g. fast and scalable run time) as compared to the PacBio. Sequencing with the Oxford Nanopores platform can also result in very long read lengths, although in contrast to PacBio where the fragment of interest is analysed multiple times, the read length on the Nanopore platform equals fragment size. Although Nanopore sequencers seem to have good features for NIPT

purposes (e.g. single molecule, base detection without labels, low GC bias, multiplexing and scalable in data output), the use of this platform for NIPT has not been described yet (Buermans *et al.*, 2014; Zhang *et al.*, 2014).

Ion Torrent Technology on the other hand has both the advantage of producing short reads in a fast and cost-efficient manner, although the total amount of data per run is relatively low compared to the previously mentioned sequencing platforms. NIPT with semiconductor sequencing has indeed been described successfully (Yuan *et al.*, 2013; Liao *et al.*, 2014). With the Ion Proton and/or improved chips for the Ion Torrent that will yield a larger number of reads, it is possible to run multiple samples at once and to handle a true clinical throughput of multiple samples per week (Wang *et al.*, 2014; Jeon *et al.*, 2014). Hence, the improved Ion Torrent and especially the Ion Proton could be good alternative sequencing platforms for NIPT purposes.

Many different factors may influence the success rate of NIPT, such as the presence of (unknown) maternal chromosomal abnormalities, CPM, true fetal mosaicism (TFM) or a vanishing twin (Bianchi *et al.*, 2012; Hall *et al.*, 2013; Wang *et al.*, 2013; Futch *et al.*, 2013; Lau *et al.*, 2014). As mentioned previously fetal fraction is one of the key factors for successful NIPT and NIPD. Fetal fraction is determined with Real-Time PCR or is calculated from NGS data. With Real-Time PCR it is difficult to determine fetal fraction when the percentage is very low. If this is the case, the data are not robust since measurement is performed in only a few copies of the fetal genome. When calculating fetal fraction from the actual sequencing data by using Y chromosomal reads, one should take into account that there is also a small percentage of reads from maternal origin that can be incorrectly assigned to the Y chromosome (Chiu *et al.*, 2011a; Hudecova *et al.*, 2014). Fetus specific Y chromosomal reads need to be deduced from the maternal background. The downside of these two methods is that they are sex dependent. In a genome wide or a targeted NGS approach, millions of reads are produced from the cfDNA present in the sample. It is generally accepted that for NIPT ~10 million unique mappable reads are required for reliable analysis. Some studies additionally stress out that for such counting-based technologies a fetal fraction of at least 4% is required for analysis (Ehrich *et al.*, 2011; Palomaki *et al.*, 2011; Palomaki *et al.*, 2012). Low fetal fraction may result in sample rejection or incorrect outcome. When the percentage of fetal sequences in a sample is too low as compared to the maternal background sequences, the relative contribution of a third chromosome in case of fetal trisomy is too low for this sample to be distinguished from the euploid foetus.

As previously described for fetal markers, both the presence and in this case quantification of cffDNA should preferably be performed with a sex independent method. The DANSR assay is an example of a NGS based sex-independent approach to determine fetal fraction. Here a set of 192 SNPs-containing loci on chromosome 1-12 was assessed for fetal trisomy detection (Sparks *et al.*, 2012a; Ashoor *et al.*, 2013; Brar *et al.*, 2013). This SNP based approach is a targeted approach, since analysis involves a selection of autosomes. However, for WGS such a SNP based targeted approach is not cost efficient at this moment, since it requires a much higher horizontal and especially vertical coverage of SNPs to be analysed. The study of Lo *et al.*, in which they show by using WGS that the complete fetal genome is present in maternal plasma, already indicated that extremely large amounts of reads are required in WGS to have sufficient coverage of a given SNP (Lo *et al.*, 2010). At present, sufficient sequencing depth cannot be accomplished without extremely high costs and is therefore not feasible yet. Another method to determine fetal fraction in a sex independent way that could be of interest for NIPT was described by Yu *et al.* They describe the analysis of library fragment size by

determining a ratio between short fetal and longer maternal fragments (Yu *et al.*, 2014). The advantage of this method is that it does not require additional experiments since library fragment size determination is widely used in standard protocols and is therefore easily applied for this purpose as well.

Monogenic disorders

For the detection of monogenic disorders, at present a genome wide approach by use of NGS is not cost-efficient. To ensure that a given mutation in the fetus will be covered sufficiently to be detected in maternal background, this would require high vertical coverage of the region of interest (Lo *et al.*, 2010). For NIPT only ~10 million unique mappable reads are required for analysis. In the paper by Lo *et al.* where a genome wide approach was used for the detection of a monogenic disorder, almost 4 billion reads were produced, equivalent to an average 65-fold coverage of the human genome (Lo *et al.*, 2010). Even then, the fetal-specific read sequencing depth of a SNP ranged from only 1-8 reads per SNP. The detection of monogenic disorders by use of a genome wide approach would ultimately lead to much higher sequencing costs since the amount of data necessary for analysis is many times higher as compared to NIPT. A whole exome sequencing (WES) approach would therefore be a better alternative for NIPD of monogenic disorders. Instead of focussing on the entire genome, only the protein-coding genes are targeted, representing ~1% of the human genome. However, WES also has its limitations for NIPD. Even though mutations in exons are more likely to have severe consequences, intronic mutations can occur and these intronic mutations will not be targeted by use of exome sequencing. Targeted gene or disease specific approaches may be custom designed for some genetic conditions, targeting genes with causative mutations. However, certain regions of the genome or exome may be difficult to target, such as GC rich regions and repeats.

Instead of using NGS, in **chapter 5 and 6** we have described alternative methods for the detection of monogenic disorders. Because of the wide variety in genetic defects, (e.g. point mutation or repeat expansion) each defect may require a different approach. In **chapter 5** we have pursued the option of blocking the maternal background in maternal plasma to enhance the detection of the fetal paternally inherited mutation. We showed that the detection of a mutation in the fetus can be enhanced by blocking the maternal background by use of complementary locked nucleic acid (LNA) probes in high resolution melting curve analysis (HR-MCA). These short LNA probes were designed to perfectly match maternal wild-type (WT) sequences. Binding of these LNA probes subsequently resulted in prevention of PCR amplification of these maternal sequences. Moreover preferential amplification and enhanced detection of the paternally inherited allele is enabled. This method can only be applied in case of PCR based targeted detection of single mutations and possibly small insertions or duplications. LNA probes are short, only around 12 nt in size. The target mutation should therefore not exceed this size. Beside HR-MCA, the use of LNA probes in Real-Time PCR could also be an attractive application in NIPD, to block maternal sequences and monitor the mutant allele in real-time. LNA probes will only bind with high affinity to complementary sequences. Because of the LNA modifications, the probe will not interfere in the PCR itself by acting as a potential primer or target. When LNA probes are used for blocking purposes, it remains essential that the parental genotypes are known, because there are highly polymorphic regions present in the human genome. In these regions the chance is high that there is a SNP present in the maternal genome at a position covered by the LNA probe. Hence, the LNA probe will not bind and consequently no inhibition of the amplification of unwanted sequences will occur. Design of

multiple LNA probes each containing one of the SNP variants could solve this problem.

As described previously for NIPT, all PCR based techniques are influenced by the GC content of the template. Therefore, GC rich areas of the human genome (e.g. promoter areas) are more difficult to amplify or target. The same holds for AT rich areas. First attempts to design HR-MCA detection and LNA probes for a familial mutation in a GC rich area in exon 4 of the *NOTCH3* gene resulted in non-functional primers and a detection probe with a very high melting temperature (Tm) (unpublished data). A difference between Tm of the LNA probe and Tm of the primers is required to facilitate LNA binding and thus blocking prior to primer binding and extension of the template sequences. The restriction of the amplicon size in combination with high GC content of the template was the main reason for failure in setting up an HR-MCA assay for this particular mutation. Increasing amplicon size could facilitate in the design of primers with a lower Tm when this genomic region is less GC rich. A strategy of blocking maternal background sequences is not necessary when the focus is on detecting polymorphic paternally inherited (extended) CAG repeats such as in Huntington disease (HD). Generally the maternal repeats are in the normal range. Size selection or selective blocking of shorter repeats in PCR cannot be accomplished with a blocking approach as was described for the detection of paternally inherited mutations. Because LNA probes target a specific sequences of ~12 nt, they are not able to distinguish between shorter or longer repeats, and in theory would block every repeat.

The size of fragmented cffDNA is also a restriction in PCR-based assays in general, for example in amplicon design. The large majority of cffDNA fragments have been shown to be on average <150 bp in length (Fan *et al.*, 2010; Lo *et al.*, 2010). This implies that these size constraints should be taken into account when designing NIPD assays to detect a given fetal target. In case a disease is caused by a repeat expansion, such as the CAG expansions in HD, the detection becomes more complex. These repeats vary greatly in length and are known to be unstable upon transmission. Therefore, repeat size is not predictable in advance and repeats may expand into a repeat size that exceeds the fragment size of cffDNA. When analysing a point mutation or small insertion or deletion, the constraint can be accommodated quite easily by designing small amplicons of ~100 bp to ensure efficient amplification of the target sequence. Fetal sequences >150 bp are also present in maternal plasma, although in minority. As shown in the original paper on the discovery of cffDNA in maternal plasma, Y chromosome specific primers designed to amplify a sequence (*DYS14*) with an amplicon size of 198 bp were used (Lo *et al.*, 1990; Lo *et al.*, 1997). This is larger than the average amplicon size for cffDNA of ~143 bp that was reported a few years later by the same group (Lo *et al.*, 2010). These results show that there is quite a range in cffDNA size. For PCR-based approaches for mutation detection, preferably amplicon sizes should be used that do not exceed the average reported size of ~143 bp.

In **chapter 6** we have described the application of NIPD for HD by directly measuring the fetal paternally inherited CAG repeat from the Huntingtin (*HTT*) gene in maternal plasma. This has been accomplished through a PCR based approach and subsequent fragment analysis of total cfDNA. For this approach to be successful for NIPD, it requires a PCR protocol that is optimized for low DNA input. Also, the Taq polymerase should prevent preferential amplification of smaller fragment as much as possible. Even though preferential amplification could not be completely prevented, the *Taq* polymerase used in PCR for NIPD of HD shows that longer fragments are being amplified in sufficient amounts to be detected in fragment analysis. This is absolutely essential for detecting the fetal expanded paternally inherited repeats in maternal plasma. When directly targeting fetal paternally inherited CAG repeats in Huntington disease,

we showed that we could detect a trinucleotide repeat up to 53 and even 70 CAG repeats, representing cffDNA fragments with a size of >170 bp and larger. As mentioned previously, fragments of this size have indeed been reported, though with a very low occurrence, which makes the detection of such rare long fragments even more challenging (Chan *et al.*, 2004). Besides restrictions in the maximum size of repeats in the expanded HD range, our results show that there are restrictions in detecting the paternally inherited repeats in the normal range. No paternally inherited repeat could be identified in approximately 50% of the cases presented, because the paternally inherited repeat coincided with the maternal repeat pattern. When NIPD results are inconclusive due to absent or low levels of cffDNA, prospective parents are given the option to provide a new blood sample for NIPD later on in gestation to retest for the paternally inherited fetal repeat. Nevertheless, in all cases where no paternal contribution in maternal plasma could be detected or in cases with inconclusive results, an invasive procedure to confirm the fetal genotype on fetal gDNA is indicated.

Fragment size analysis is at present the best strategy for NIPD for HD. Detection of large repeats for NIPD using other technologies such as NGS is currently difficult, if not impossible. CAG repeats that encode for long glutamine (Q) stretches are present all through the human genome. To date, a total of nine polyQ diseases have been reported (Fan *et al.*, 2014). A NGS mediated approach for repeat determination and/ or detection for NIPD of HD would require paired-end sequencing of a specified and targeted repeat, in this case the *HTT* gene. Although the repetitive sequence of a repeat is simple, errors in NGS data may interfere with correct data analysis. Moreover, since CAG repeat are common throughout the genome, bioinformatics will have to focus on HTT specific CAG repeats to determine repeat size confidently.

Concluding remarks

Noninvasive prenatal diagnostics and testing have strongly improved in time and more and more applications have become available. Since the entire fetal genome is present in maternal plasma, it is to be expected that in the future the large majority if not all of prenatal diagnostics is preferably be performed using cffDNA. Whether or not this will always be performed in a universal and genome wide approach using WGS has to be elucidated. Currently, there are still limitations in using WGS for some prenatal requests. With WGS large amounts of data can be generated, which can be useful in detecting known but also unknown variants or de novo mutations in fetuses without a prior history or predisposition of a familial mutation. However, shotgun sequencing is not mutation specific and is currently not always cost-efficient for use in the detection of these specific mutations.

The majority of prenatal requests are for fetal aneuploidy detection. In the Netherlands, during a two year national implementation study, NIPT is currently available for high risk pregnant women. It is to be expected that NIPT will soon become available for low risk pregnant women as well. In addition to this, in the near future it is desirable for genetic laboratories to have several NIPD alternatives available for prospective prenatal requests. As mentioned previously and as shown in this thesis, no universal approach, particularly for NIPD, is available yet. For each novel application, one should clearly consider the best approach in detecting the genetic defect using NIPD. Moreover, some applications (e.g. detection of point mutations) do not have to be restricted only to couples where the father is carrier of the mutant allele.

In conclusion, we show several novel applications for the use of cffDNA for NIPD and NIPT. We also show that each application at present may require a different approach. In the near future, it is to be expected that more noninvasive alternatives for a wider variety of genetic anomalies will become available for prenatal diagnostics early in gestation.