# Chapter 3

# Single Molecule Sequencing of Free DNA from Maternal Plasma for Noninvasive Trisomy 21 Detection

# Chapter 3: Single Molecule Sequencing of Free DNA from Maternal Plasma for Noninvasive Trisomy 21 Detection

Jessica van den Oever

Sahila Balkassmi

Joanne Verweij

Maarten van Iterson

Phebe Adama van Scheltema

Dick Oepkes

Jan van Lith

Mariëtte Hoffer

Johan den Dunnen

Bert Bakker

Elles Boon

Clin Chem, 2012 Apr;58(4):699-706. doi: 10.1373/clinchem.2011.174698.

## *Abstract*

**Background:** Noninvasive fetal aneuploidy detection using free DNA from maternal plasma has recently been shown to be achievable by whole genome shotgun sequencing. The high-throughput Next Generation Sequencing platforms previously tested use a PCR step during sample preparation, which results in amplification bias in GC rich areas of the human genome. To eliminate this bias, and thereby experimental noise, we have used single molecule sequencing as an alternative method.

**Methods:** For noninvasive trisomy 21 detection, single molecule sequencing was performed on the Helicos platform using free DNA isolated from maternal plasma from 9 weeks of gestation onwards. Relative sequence tag density ratios were calculated and results were directly compared to the previously described Illumina GAII platform.

**Results:** Sequence data generated without an amplification step show no GC-bias. Therefore, using single molecule sequencing all trisomy 21 fetuses could be distinguished more clearly from euploid fetuses.

**Conclusion:** This study shows for the first time that single molecule sequencing is an attractive and easy to use alternative for reliable noninvasive fetal aneuploidy detection in diagnostics. Using this approach, previously described experimental noise associated with PCR amplification, such as GC bias, can be overcome.

## *Introduction*

Trisomy 21 (T21) is the most common chromosomal abnormality in live-born children. The diagnosis can be made early in pregnancy using invasive testing (e.g. chorionic villus sampling (CVS) or amniocentesis). These invasive procedures however, are associated with a risk of miscarriage. Therefore, these tests are commonly only offered to women at increased risk for fetal trisomy. Risk assessment used to be based on maternal age. More recently, this was refined by adding serum markers for trisomy and ultrasound measurement of the fetal nuchal translucency (Driscoll *et al.*, 2009). Current screening programs have detection rates for T21 of around 80% with a false positive rate of 5%, meaning that one in every 20 women screened is offered invasive testing with its inherent risks, while carrying a healthy fetus (Wapner *et al.*, 2003; Mujezinovic *et al.*, 2007).

The discovery of cell-free fetal (cff) DNA and RNA in maternal plasma opened possibilities for noninvasive prenatal diagnosis (NIPD) (Lo *et al.*, 1997). Although cffRNA has been used for noninvasive T21 detection (Lo *et al.*, 2007b; Picchiassi *et al.*, 2010; Tsui *et al.*, 2010; Deng *et al.*, 2011), the majority of approaches use cffDNA for NIPD of T21. In the first trimester, the percentage of cffDNA in maternal plasma is on average 1-10% and differs quite extensively in range depending on gestational age and between individuals (Go *et al.*, 2010; Chiu *et al.*, 2011b; Lo *et al.*, 1998; Lun *et al.*, 2008a; Sikora *et al.*, 2010; Hahn *et al.*, 2011). Therefore, it remains challenging to detect fetal sequences in a large pool of maternal DNA. Previously, several papers have shown that noninvasive T21 detection is possible by using single nucleotide polymorphisms (SNPs) (Dhallan *et al.*, 2007; Ghanta *et al.*, 2010) and epigenetics (Old *et al.*, 2007; Chim *et al.*, 2008; Tong *et al.*, 2010b; Papageorgiou *et al.*, 2011) although these methods have a number of limitations.

In 2008, noninvasive T21 detection by Next Generation Sequencing (NGS) was introduced (Fan *et al.*, 2008; Chiu *et al.*, 2008), opening a whole new way of analysis. No longer only fetal specific sequences were analyzed, but all free DNA in plasma, from both fetal and maternal origin, is sequenced with this technique. Two recent papers confirmed the potential value of NGS for noninvasive fetal T21 detection in multiplexed plasma DNA samples in a clinical setting (Ehrich *et al.*, 2011; Chiu *et al.*, 2011a). Both the Illumina Genome Analyzer (GA) II (Fan *et al.*, 2008; Chiu *et al.*, 2008; Fan *et al.*, 2010; Ehrich *et al.*, 2011; Chiu *et al.*, 2011a) and the SOLiD platform (Chiu *et al.*, 2010) have been used for noninvasive T21 detection by NGS. These platforms use amplification steps by polymerase chain reaction (PCR) which are known to introduce preferential amplification of sequences depending on different GC content (Fan *et al.*, 2008; Chiang *et al.*, 2009).

In the present study, we have tested single molecule sequencing (tSMS, Helicos Heliscope™ Single Molecule Sequencer) for noninvasive T21 detection. The Helicos platform utilizes visual imaging across the flow cell for direct DNA measurement by recording the incorporation of fluorescently labeled nucleotides (Gupta, 2008; Milos, 2009). The use of single molecule sequencing has been described previously (Harris *et al.*, 2008) and this technique should largely overcome the limitations associated with PCR amplification and bias as mentioned above. Although the sequencing time on the Helicos platform is longer compared to the Illumina platform (4 days respectively 2 days), Helicos sample preparation is simple, 3 times faster (1 day compared to 3 days) and therefore relatively cheap. Furthermore, this method requires low amounts of DNA, which could be of special interest early in gestation.

Here, we present a comparison of the application of single molecule sequencing for noninvasive T21 detection using cffDNA from maternal plasma to the previously described PCR-based Illumina NGS platform.

## *Materials and Methods*

### *Subjects*

Pregnant women undergoing prenatal diagnosis were recruited at the Department of Obstetrics of the Leiden University Medical Center (LUMC), Leiden, The Netherlands. Informed consent was obtained and this study was approved by the Institution's Medical Ethics Committee.

### *Sample Processing and Isolation*

Maternal peripheral blood samples (10-20 mL) were collected in EDTA coated tubes at the LUMC and were processed within 24 hrs after collection. All blood samples were drawn at a median gestational age of 12 +2 weeks (range 9 +3 to 16 +6 wks). Preferably blood samples were drawn before an invasive procedure, if this was not possible samples were drawn at least 5 days after the invasive procedure to minimize any disturbance with fetal material due to this procedure.

Blood was centrifuged at 1200g (without brake) for 10 min at room temperature. Plasma was transferred to 15 mL micro centrifuge tubes and centrifuged at 2400g for 20 min (with brake) at room temperature to remove residual cells. Cell-free plasma was divided into 800 µL aliquots and stored at -80$^{o}$C until further processing.

Because both sequencing platforms require different amounts of input DNA, cell-free DNA was isolated from plasma with the EZ1 Virus Mini Kit v2.0 on the EZ1 Advanced (QIAGEN, Venlo, The Netherlands; www.qiagen.com) for Helicos sample preparation or manually with the QiaAmp MinElute Virus Spin Kit (QIAGEN) for Illumina sample preparation according to the manufacturer's instructions.

To verify fetal gender and to measure the total quantity of cell-free DNA, we respectively performed a pyrophosphorylation-activated polymerization assay on the Y chromosome (Y-PAP) and a Real-Time Taqman PCR assay on *CCR5* for quality control purposes as described previously (Boon *et al.*, 2007). In addition, for male fetuses we estimated the percentage of cffDNA based on sequencing data of chromosome X (Fan *et al.*, 2008) and by Real-Time Taqman PCR assay on *SRY*, for which we used a standard curve from male genomic DNA to determine the range of cffDNA percentages in maternal plasma. Percentages were estimated by dividing the amount of *SRY* (pg/µL) by the maternal fraction of *CCR5* from 1 allele (*SRY*/(0.5**CCR5*$_{total}$ – *SRY*)), taken into account that the PCR efficiency of both genes is similar.

### *Library preparation and sequencing*

A total of 24 plasma samples was included in this retrospective study, containing 20 samples from singleton pregnancies, of which 11 cases (5 female and 6 male fetuses) of T21, 9 cases of disomy (D21) pregnancies (1 female and 8 male fetuses) and 4 plasma control samples from anonymous adult male blood donors. All samples were de-identified to the investigators before sample preparation and data analysis. These results were not revealed to the investigators until after data analysis. Material from the invasive procedure was sent to the

Cytogenetics Lab for full karyotyping. Fetal gender was confirmed by karyotype or after birth.

All cell-free plasma DNA samples were sequenced on both the Helicos (Helicos BioSciences Corporation, Cambridge, MA, USA, www.helicosbio.com) and the Illumina (Illumina Inc., San Diego, CA, USA, www.illumina.com) GA II platform. Owing to the relatively short length (Fan *et al.*, 2008) and fragmented nature of free DNA in plasma, no additional shearing step was performed during library preparation.

Helicos sample preparation was performed according to the manufacturer's ChIP-Seq Direct Tailing Procedure with an input of 400 µL plasma for DNA isolation with the EZ1 (QIAGEN) and the maximum amount of input for tailing. As a quality control, size of the fragments and template size distribution were determined by running a High Sensitivity DNA chip on the Agilent Technologies 2100 Bioanalyzer. A standard 120-cycle run was performed on the Heliscope™ Single Molecule Sequencer, which resulted in an average read length of 35 nucleotides.

Illumina sample preparation was performed according to the manufacturer's ChIP-Seq protocol with an input of 1600 µL plasma per sample per column for manual DNA isolation and a maximum amount of input for this protocol. Sixteen out of 24 samples were sequenced in a duplex assay (T21 n=7, D21 n=7 and male plasma controls n=2). For this, unique synthetic 6 nucleotide barcodes (indexes) were used. The barcode was ligated to the plasma DNA molecule prior to the PCR amplification step. Indexed samples were additionally purified on a 3% TAE Agarose gel prior to the quality control run on the Bioanalyzer as mentioned above. A 36-cycle run was performed on the Illumina GA II.

### *Data analysis*

Helicos sequencing data were analyzed with the Helicos Helisphere resequencing pipeline using default settings. Data were aligned against hg19 and gaps and repeats were filtered out. Filtered data were sorted and binned per 50 kb.

Illumina raw data from duplexed samples were pre-analyzed by splitting the data per indexed barcode with in-house Linux command lines. Sequencing data were analyzed with NextGENe software (SoftGenetics, State College, PA, USA, www.softgenetics.com). Data were mapped to the annotated Human Genome GFCh37-dbSNP 131(4/14/2010) (hg19) for Illumina data compatible with NextGENe software. Expression reports per 50 kb were created. Only unique reads with at most 1 mismatch, which could be aligned to the reference genome, were used for calculations.

For all samples (both T21 and D21) used for noninvasive fetal trisomy 21 detection in maternal plasma, ratios of relative sequence tag density (RSTD) were calculated. First, for each sample the total number of reads was calculated per chromosome, by summing the read counts of all 50 kb bins belonging to a particular chromosome. Second, for each sample, the total summed number of reads was normalized by the median value of the autosomes. Finally, ratios of RSTD were calculated by dividing these normalized values by the averaged normalized value of the disomy samples (Fan *et al.*, 2008) or, in addition, by the normalized average of male plasma control samples. As the data were obtained by two separate runs for both sequencing technologies, ratios were determined for each run separately.

### *Statistical analysis*

Statistical analysis was conducted with PASW Statistics version 17.0 (SPSS Inc., Chicago IL, USA, www.spss.com), Prism 5 (version 5.00, GraphPad Software, Inc. La Jolla CA, USA, www.graphpad.com) and R version 2.13 (R Development Core Team (2011)). R: A language environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN

**Table 1: Overview of included maternal plasma samples.**

| *Sample number* | *Karyotype* | *Maternal Age* | *Indication* | *AD Blood* | *CVS/ Amnio* | *AD Procedure* |
|---|---|---|---|---|---|---|
| **1** | 46,XY | 35 | Family history of mental retardation/ ICSI pregnancy | 12 +2 | Amnio | 16 +1 |
| **2** | 46,XY | 39 | Advanced maternal age | 9 +6 | Amnio | 17 +0 |
| **3** | 47,XX,+21 | 41 | Ultrasound abnormality: Hygroma colli and generalized oedema | 15 +3 | CVS | 14 +2 |
| **4** | 47,XY,+21 | 43 | Advanced maternal age | 10 +2 | Amnio | 16 +0 |
| **5** | 46,XX | 40 | Advanced maternal age | 10 +1 | Amnio | 16 +0 |
| **6** | 46,XY | 38 | Advanced maternal age | 11 +2 | Amnio | 17 +1 |
| **7** | Male | 39 | Increased NT, advanced maternal age | 9 +3 | - | - |
| **8** | Male | 37 | Increased NT, advanced maternal age | 13 +2 | - | - |
| **9** | 47,XY,+21 | 34 | Ultrasound abnormality | 13 +2 | CVS | 12 +0 |
| **10** | Male | 38 | Increased NT, advanced maternal age | 11 +0 | - | - |
| **11** | 47,XX+21 | 41 | Advanced maternal age | 16 +6 | Amnio | 16 +1 |
| **12** | 47,XY,+21 | 41 | Advanced maternal age | 14 +4 | CVS | 12 +3 |
| **13** | 47,XX,+21 | 41 | Ultrasound abnormality: Hydrothorax, cystic hygroma | 11 +2 | CVS | 11 +3 |
| **14** | 46,XY | 40 | Advanced maternal age | 11 +5 | CVS | 11 +5 |
| **15** | 46,XY | 39 | Advanced maternal age | 16 +6 | Amnio | 16 +6 |
| **16** | 47,XY,+21 | 38 | Ultrasound abnormality | 12 +2 | CVS | 12 +2 |
| **17** | 47,XY,+21 | 39 | Ultrasound abnormality: Hygroma colli | 12 +1 | CVS | 12 +1 |
| **18** | 47,XX,+21 | 36 | Ultrasound abnormality: Hygroma colli | 14 +2 | CVS | 14 +2 |
| **19** | 47,XY,+21 | 35 | Increased NT | 12 +4 | CVS | 12 +4 |
| **20** | 47,XY,+21 | 34 | Ultrasound abnormality: Hygroma colli | 13 +1 | CVS | 13 +1 |

**Table 1:** *AD Blood*: Gestational age at the time of blood collection depicted as weeks +days, *CVS*: Chorionic Villus Sampling, *Amnio*: Amniocentesis, *AD Procedure*: Gestational age at the time of the invasive procedure depicted as weeks +days, *ICSI:* Intracytoplasmic Sperm Injection, *Increased NT*: Increased nuchal translucency thickness.

3-900051-07-0, www.R-project.org). Differences between the numbers of uniquely mapped reads between groups were determined by independent samples T-test. Correlation between the number of reads and RSTD were determined by non-parametric Spearman correlation. P values of less than 0.05 are considered statistically significant.

## *Results*



### *Included samples*

A total of 20 maternal plasma samples were included in this study and were taken at a median gestational age of 12 +2 weeks (range 9 +3 to 16 +6 weeks). In 4 out of 20 cases blood samples were drawn after the invasive procedure (on average > 1 week afterwards). No correlation between the time of sampling (before or after the invasive procedure) and the ratios was observed. All details on the included samples are depicted in Table 1. For the noninvasive detection of fetal T21, DNA isolated from 20 maternal plasma samples and 4 anonymous male plasma controls were sequenced on both the Helicos and the Illumina GA II platform. One D21 sample, that did pass the quality controls prior to sequencing, failed the quality controls after sequencing for both platforms. For this sample hardly any reads were obtained for the Helicos platform and sequencing results from the Illumina platform showed preferential amplification of only a few regions. This sample was therefore excluded for further analysis.

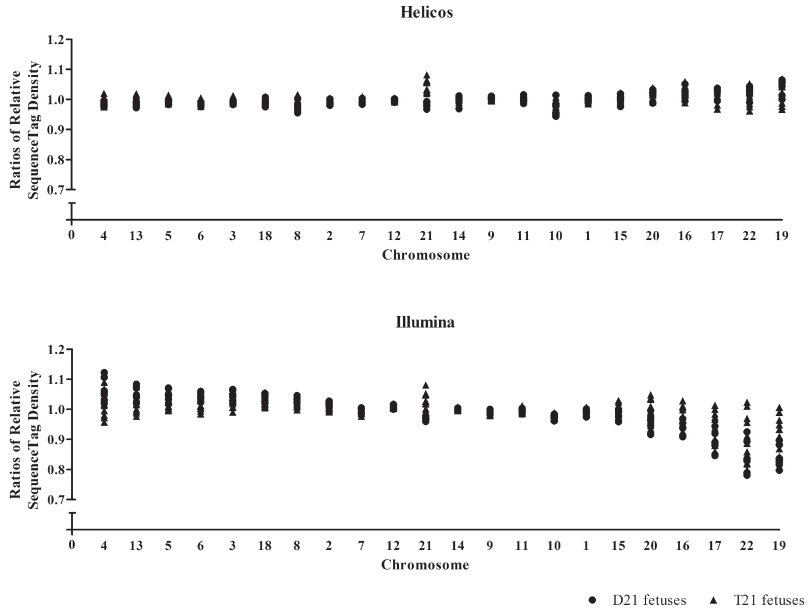



**Figure 1: Ratios of normalized relative sequence tag density (RSTD) from all autosomes.**
Ratios are calculated against averaged normalized read counts from male plasma controls. Data are shown for each Next Generation Sequencing platform (T21 n=11 and D21 n=8). Chromosomes are ordered by increasing GC content. Upper panel: Helicos, Lower panel: Illumina GA II.

## *Sample statistics*

For each NGS platform, the mean number of raw reads, the percentage of filtered reads and the mean and median number of uniquely mapped reads are depicted in Table 2. For Helicos, our data show one D21 sample with the overall lowest amount of reads, to have the lowest RSTD ratio, but overall we observed no correlation between RSTD ratio and the amount of uniquely mapped reads for both platforms (Helicos, Spearman r = -0.088, 95%CI [-0.532-0.394], P = 0.7210 and Illumina, Spearman r = -0.232, 95% CI [-0.629, 0.263], P = 0.3401). Furthermore, the number of uniquely mapped reads between T21 and D21 was similar (Helicos P = 0.128 and Illumina P = 0.810). When looking at the duplexed Illumina samples (n=16), no bias in read counts was observed towards any specific barcode after splitting (P= 0.9551).

The percentage of cffDNA in maternal plasma was calculated using 2 different methods. When using the method based on Illumina sequence data from chromosome X by the group of Fan *et al*. (Fan *et al.*, 2008), we estimated the percentage of fetal DNA for male pregnancies (n=6) to be on average ~7% (range 1-18%). Concordant results were obtained by Real-Time PCR on the *SRY* gene (average ~9%, range 3-18%).

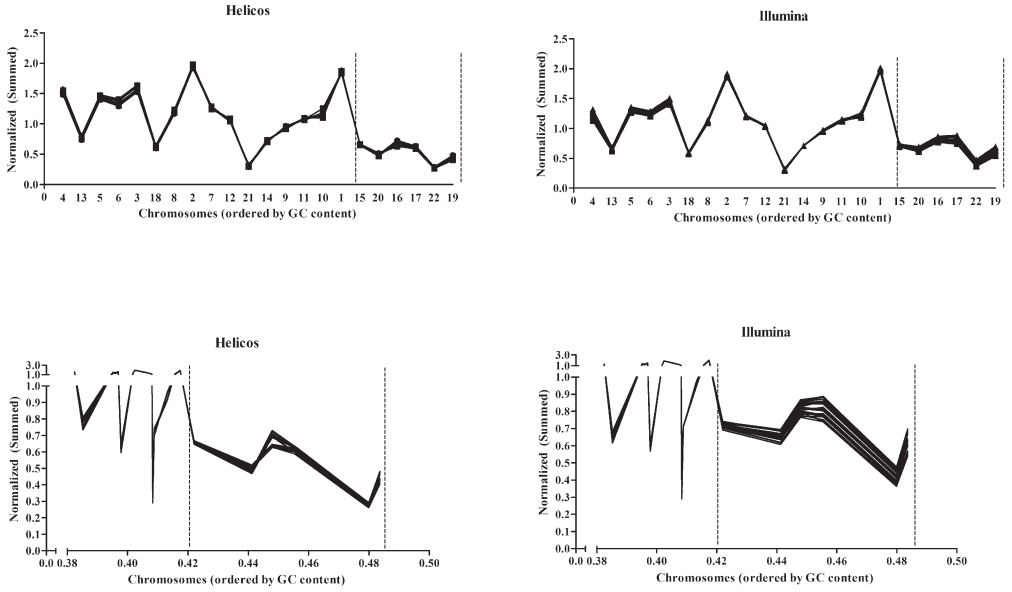



**Figure 2: Normalized total number of reads per chromosome against GC content.**
Normalized reads are shown in order of GC content per chromosome (upper panel) and by GC percentage (lower panel) for both the Helicos (left) and Illumina (right) platform. Chromosomes subject to possible GC bias because of high GC content are depicted within dashed lines (upper and lower panel).

## *Noninvasive T21 detection*

For the detection of noninvasive fetal T21, RSTD ratios for all 19 maternal plasma samples are shown per chromosome for each NGS platform (Fig. 1). The autosomes were ordered by increasing GC content (Fan *et al.*, 2008). The overall distribution of reads across the genome

is similar between both platforms and seems independent of GC content (data not shown). However, our data show a clear difference in read coverage between platforms. For Helicos, the RSTD ratios for all chromosomes (Fig. 1), the normalized total number of reads per chromosome (Fig. 2) and the average amount of reads per bin (Fig. 3) were quite uniform between samples and virtually independent of GC content of the chromosome, while as reported before (Fan *et al.*, 2008), Illumina results showed increased read density in GC rich areas of the genome (Fig. 1-3).

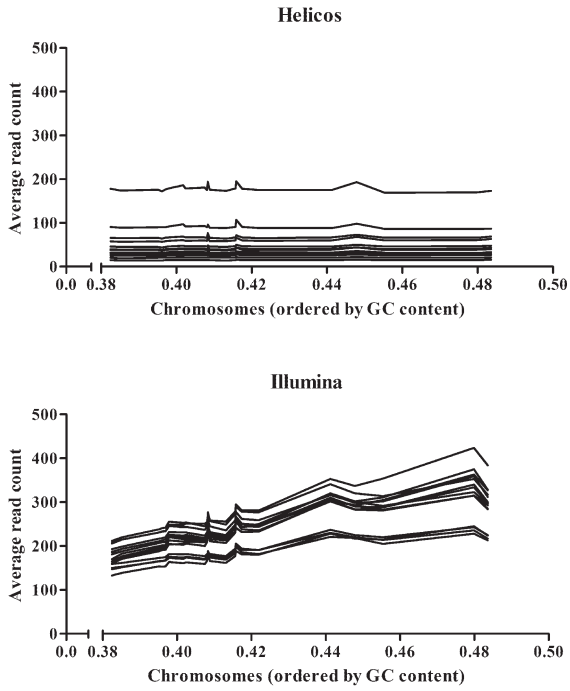



**Figure 3. Average read count per 50 kb bin against GC content.**
For both platforms, the average number of reads per bin was determined by dividing the summed total number of reads by the number of bins. Chromosomes are ordered by GC content.

Our data show RSTD ratios for T21 samples in a range of 1.04-1.11 for Helicos and a range of 1.03-1.12 for Illumina. For D21 samples we obtained RSTD ratios from 0.98-1.01 and 0.99-1.01 respectively (Fig. 4). Our data show a clear distinction between plasma samples from women carrying a T21 fetus and woman carrying a D21 fetus for both platforms when looking at the overrepresentation of the affected chromosome (Fig. 4). All maternal plasma samples of women carrying a fetus with Down syndrome were correctly classified as T21 (n=11). In addition, all euploid samples (n=8) were correctly identified as D21, resulting in a sensitivity and specificity of both 100% (95% CI [87.0-100]). When constructing a 99% confidence interval of the distribution of RSTD from all D21 samples, all T21 samples lie outside the upper boundary of 1.01 and all D21 samples on or below this boundary. Overall, we show that noninvasive detection of T21 can be performed on both NGS platforms, although Helicos results show a better distinction between T21 and D21 samples (Fig. 4).

**Table 2: Overview of mean and median number of uniquely mapped reads[A].**

| *Platform* | *Mean # of raw reads* | *% of filtered reads* | *Mean # of alligned reads* | *Median # of alligned reads* |
|---|---|---|---|---|
| Helicos | $1.06 \times 10^7$ ($0.46 \times 10^7$) | 35.09 (1.52) | $4.65 \times 10^6$ ($3.6 \times 10^6$) | $2.65 \times 10^6$ |
| Illumina | $2.47 \times 10^7$ ($0.38 \times 10^7$)[B] | 76.07 (5.14)[B] | $1.26 \times 10^7$ ($0.40 \times 10^7$) | $1.26 \times 10^7$ |

**Table 2:** [A] Results are indicated by Next Generation Sequencing Platform. Data for each platform are represented as mean (SD) for n=23 samples. [B]For the Illumina platform, the mean number of raw reads and % filtered reads is depicted for the duplexed samples n=15 (T21 n=7, D21 n=6, male plasma control n=2).

## *Calculation methods*

We have based our calculations on the method of Fan *et al.* (Fan *et al.*, 2008), which uses read counts to calculate ratios of RSTD. Samples can be normalized against averaged normalized RSTD from both adult male plasma controls (Fig. 1) or disomy samples (Fig. 4). Our results look similar when applying either one of these methods to data from both NGS platforms. Recently, a new calculation method for the detection of fetal chromosomal abnormalities was published by the group of Sehnert *et al.* (Sehnert *et al.*, 2011). With this method, samples can be classified as affected (i.e. aneuploid for that chromosome) or unaffected by calculating a normalized chromosome value (NCV) using data from a previously analyzed training set consisting of unaffected samples (i.e. maternal plasma samples from women carrying a euploid fetus). When applying this new calculation method to our Illumina data, all Illumina samples were again correctly classified as either T21 or D21, within the criteria as described (See Supplemental Figure S1) (Sehnert *et al.*, 2011). Since these criteria are only determined for Illumina data, they are not applicable on our Helicos results and thus first need to be established.

a

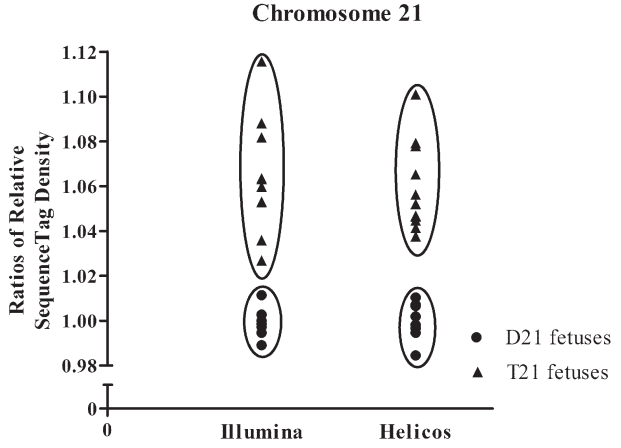



**Figure 4. Ratios of normalized relative sequence tag density (RSTD) from chromosome 21.**
Ratios are calculated against averaged normalized read counts from disomy samples. Data are shown for each Next Generation Sequencing platform (T21 n = 11 and D21 n = 8).

## *Discussion*

Noninvasive fetal aneuploidy detection using free DNA from maternal plasma has evolved dramatically the past few years with the introduction of NGS. The majority of studies use the Illumina GA II platform for whole genome shotgun detection of T21. Data obtained in these studies have shown that limitations due to low percentages of cffDNA in maternal plasma, no longer seem to be a major problem. However, the Illumina platform is PCR-based and the amplification step could initiate several negative side effects, such as read density bias in GC rich areas of the genome.

In this study, we show successful fetal T21 detection using free DNA from maternal plasma by single molecule sequencing on the Helicos platform and compared it to the Illumina GA II platform (Fan *et al.*, 2008; Chiu *et al.*, 2008; Ehrich *et al.*, 2011; Chiu *et al.*, 2011a). For Illumina, we could confirm previously described findings (Fan *et al.*, 2008). Moreover, we demonstrate a more distinct separation between T21 and D21 samples in Helicos data versus Illumina. We show that as early as 9 +3 wks of gestational age, cffDNA samples from maternal plasma can be classified correctly with high sensitivity and specificity. Because for single molecule sequencing only small amounts of free DNA are required as input for sample preparation and direct sequencing is performed, we hypothesize that this method might therefore be more suitable for early noninvasive aneuploidy detection.

Also, our study confirms that data obtained on the Helicos platform is not biased in GC rich areas, leading to an increased accuracy of analysis. Previously, a strong correlation between GC rich areas and read coverage was observed on the Illumina platform, with increased number of reads in areas containing elevated GC content (Dohm *et al.*, 2008; Hillier *et al.*, 2008; Fan *et al.*, 2008). There has been discussion whether this is a biological effect relating to chromatin structure or originates from PCR artifacts introduced during sample preparation, cluster formation or the sequencing process itself. Since GC bias is not observed in single molecule sequencing it is less likely that this is a true biological effect or be due to the sequencing process. We therefore hypothesize that it is introduced in the pre-amplification step for DNA enrichment or during local amplification for cluster formation on the flow cell. The exact reason, however, remains to be elucidated.

Before implementing noninvasive trisomy detection into routine diagnostics several quality controls criteria need to be determined and validated. The QUADAS criteria can be applied, which take into account the experimental bias and variation (Whiting *et al.*, 2003). Equally important are the quality controls before and during sample preparation. Since the percentage of cell free DNA in maternal plasma differs between samples and at different times of gestation (Lo *et al.*, 1998; Lun *et al.*, 2008a), it is difficult to determine the most appropriate time of gestation for testing. However, for diagnostics inclusion criteria including time of gestation need to be determined. Measurement of the amount of cffDNA and its correlation to reliable diagnosis and time of gestation needs to be studied more thoroughly in large validation studies. Before sequencing isolated free DNA, combined Real-Time PCR results on *CCR5* and *SRY* could help estimate the ratio of maternal and fetal DNA in maternal plasma, the percentage of fetal DNA and the quality of DNA as shown in our data. After sequencing, percentages of cffDNA can then be verified by using data from chromosome X as described previously (Fan *et al.*, 2008). Both methods however are limited to male pregnancies only. When encountering samples containing low percentages of fetal sequences or large amounts of contaminating maternal sequences, restrictions for the detection limit should be taken into account. For female pregnancies a sex-and polymorphism-independent method based on epi-

genetic differences could be used for quantification (Nygren *et al.*, 2010), although it needs to be established whether differences in methylation are stable and comparable between individuals to be used reliably in diagnostics.

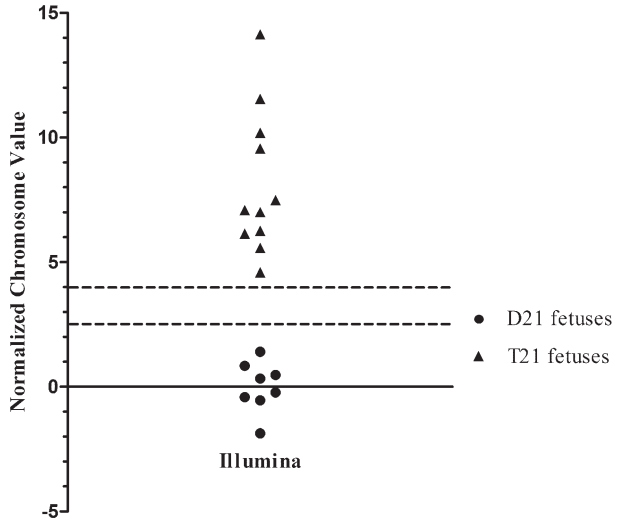
Another issue that should be taken into account are maternal copy number variations (CNVs). These can be of particular interest for the interpretation of trisomy detection using NGS data in diagnostics. Pre-determination of CNVs in the maternal genome could be an useful control in diagnostics, because these findings may influence the interpretation of data when looking at the overrepresentation of a specific chromosome, regardless the NGS platform used.

In summary, this study shows for the first time that single molecule sequencing can be a reliable and easy-to-use alternative for noninvasive T21 detection in diagnostics. By using single molecule sequencing, previously described experimental noise associated with PCR amplification, such as GC bias, can be overcome. This method is therefore not only promising for noninvasive T21 detection, but is potentially also useful for the detection of other aneuploidies.

## *Acknowledgements*

We would like to thank Jennie Verdoes for including pregnant women, Michiel van Galen for bioinformatics, Yavuz Ariyurek and Henk Buermans for technical support and BIOKÉ (The Netherlands) for NextGENe software assistance.

## *Supplemental data*




**Supplemental figure S1. Normalized Chromosome Value for Chromosome 21.**
Normalized Chromosome Value (NCV) for chromosome 21 was calculated according to the method described in Sehnert *et al*. (Sehnert *et al.,* 2011). An NCV > 4.0 was used to classify the sample as aneuploid for chromosome 21 and an NCV < 2.5 to classify a chromosome as unaffected. Samples with an NCV between 2.5 and 4.0 were classified as "no call" (dashed lines).