ORIGINAL ARTICLE

Comparing methods for fetal fraction determination and quality control of NIPT samples

Daphne M. van Beek[1†], Roy Straver[1†], Marian M. Weiss[1], Elles M. J. Boon[1,2], Karin Huijsdens-van Amsterdam[3], Cees B. M. Oudejans[4], Marcel J. T. Reinders[5] and Erik A. Sistermans[1*]

[1]Department of Clinical Genetics, VU University Medical Center Amsterdam, Amsterdam, The Netherlands
[2]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands
[4]Department of Clinical Chemistry, VU University Medical Center Amsterdam, Amsterdam, The Netherlands
[5]Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands
*Correspondence to: E. A. Sistermans. E-mail: e.sistermans@vumc.nl
†These authors contributed equally to this work.

ABSTRACT

Objective To compare available analysis methods for determining fetal fraction on single read next generation sequencing data. This is important as the performance of non-invasive prenatal testing (NIPT) procedures depends on the fraction of fetal DNA.

Methods We tested six different methods for the detection of fetal fraction in NIPT samples. The same clinically obtained data were used for all methods, allowing us to assess the effect of fetal fraction on the test result, and to investigate the use of fetal fraction for quality control.

Results We show that non-NIPT methods based on body mass index (BMI) and gestational age are unreliable predictors of fetal fraction, male pregnancy specific methods based on read counts on the Y chromosome perform consistently and the fetal sex-independent new methods SeqFF and SANEFALCON are less reliable but can be used to obtain a basic indication of fetal fraction in case of a female fetus.

Conclusion We recommend the use of a combination of methods to prevent the issue of reports on samples with insufficient fetal DNA; SANEFALCON to check for presence of fetal DNA, SeqFF for estimating the fetal fraction for a female pregnancy and any Y-based method for estimating the fetal fraction for a male pregnancy. © 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.

Funding sources: None
Conflicts of interest: None declared

INTRODUCTION

Pregnant women undergoing invasive prenatal tests have a 2–3:1000[1] risk for an iatrogenic abortion. With the introduction of non-invasive prenatal testing (NIPT), pregnant women have a safe alternative to test for fetal aberrations such as Down syndrome (trisomy 21), Patau syndrome (trisomy 13) and Edwards syndrome (trisomy 18).[2] Reliability of NIPT is high but depends on biological as well as experimental variations. It has been speculated that the fetal fraction should be at least 4%[3] to allow for reliable detection of common trisomies. Several methods to predict or determine fetal fraction have been proposed. Previous publications indicate that high body mass index (BMI) negatively affects the fraction of fetal DNA in maternal plasma,[4] although these correlations were not very strong. Due to their high variability, these measures are not being used in a clinical workflow. Dedicated laboratory tests to determine fetal fraction directly on isolated DNA have been developed but have the associated risk of adding an error prone step in the diagnostic work-up. When splitting the sample in two different lab flows, one for determination of fetal fraction and the other for the library prep, an error or sample swap in either flow might result in a mismatch between the fetal fraction and the NIPT result. Furthermore, incorrect preparation for sequencing may result in loss of the fetal DNA or poor size selection, which will not be noticed if the fetal fraction is determined on a different lab flow. These kinds of technical failures might even affect an entire run of samples. Therefore, the fetal fraction should preferably be determined from the same next generation sequencing (NGS) data that are used for the determination of the chromosomal aberrations. There are multiple methods developed for the detection of fetal fraction in cfDNA samples, but a comparison

© 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

between these methods is still lacking. For male fetuses, the easiest approach is to use the relative amount of reads that map on the Y chromosome as this is an unambiguous indication of fetal DNA. We compare two Y chromosome-based methods, the first being a new method we developed, DEFRAG. This method calculates both a normalized fraction with respect to pure male/female DNA, as well as a fraction that takes into account the uniqueness of the Y chromosome. DEFRAG was compared to a method introduced by Bayindir *et al.*,[5] to which we further refer as BAYINDIR, that uses a robust estimate of the fraction of reads mapping on the Y chromosome with respect to the autosomes. Both DEFRAG and BAYINDIR calculate the fetal fraction using two independent methods. Recently, two methods have been introduced that are capable of estimating the fetal fraction independent of fetal sex.[6,7] It is yet unclear how these methods perform in a clinical setting. We set out to evaluate and compare these methods in their ability to determine the fetal fraction from NGS-NIPT data. As this article focuses on methods that use relatively cheap short single read sequencing, a recent method that uses insert size based on paired-end sequencing[8] is excluded. Finally, we related fetal fraction to BMI and gestational age.

METHODS

DNA was isolated from 654 blood plasma samples (279 pregnancies of a female fetus and 375 pregnancies of a male fetus). Metadata on pregnancies were collected during the TRIDENT study in the Netherlands; these included BMI and gestational age.[9] All women signed an informed consent form. Permission for the study was granted by the Minister of Health (11016-118701-PG). The study was also approved by local University Medical Center Ethics Committees. DNA was sequenced using the Illumina HiSeq2500, obtaining 4 to 37 million (median 16 million) 51 BP single end reads per sample. Sequence data were demultiplexed allowing one mismatch in the index tag, mapped to GRCh37 using BWA, allowing zero mismatches and removing any read that had multiple mappable positions. For the training of DEFRAG and BAYINDIR, an additional training set of 196 samples (76 female pregnancies, 109 male pregnancies and 11 male control samples) was used.

DEFRAG

DEFRAG consists of two methods that both use the number of reads mapped to the Y chromosome to determine the fetal fraction and is thus limited to use on male pregnancy samples.

$DEFRAG_a$: Normalized fraction of reads on chromosome Y

Female pregnancy samples from the training set were used to determine the average fraction of reads mapped to chromosome Y at 0% male DNA ($\%Y_{XX\ fetus}$). The 11 male control samples were used to determine this fraction at 100% male DNA ($\%Y_{XY\ man}$). The fetal fraction in a new, unseen, sample is then determined by the fraction of reads mapped to Y with respect to these bounds:

$$DEFRAG_a = \frac{\%Y_{XY\ fetus} - \%Y_{XX\ fetus}}{\%Y_{XY\ man}} \quad (1)$$

where $\%Y_{XY\ fetus}$ refers to the percentage of reads in the sample that map to the Y chromosome.

$DEFRAG_b$: Fraction of reads uniquely mapped to chromosome Y

Even in case of a female pregnancy, reads map to some parts on the Y chromosome due to repeated regions. To increase sensitivity for male pregnancies, we calculated a fetal fraction in which we only considered those regions on Y where only rarely reads map in case of a female pregnancy. Chromosome Y is divided in 1 Mb bins and the amount of reads in each bin is determined for the samples in the training set. Bins that never received any reads across all samples, or contain at least one read in more than half of the female pregnancies, are removed (excluding regions 0–2, 9–14, 20–21 and 25–59 Mb on ChrY). This results in a Y-specific subset of bins containing reads in male pregnancies, but rarely in female pregnancy samples. These bins are corrected for GC content using a LOWESS function, which is trained for the expected number of reads given the GC fraction of a bin using the autosomes.[10] The fetal fraction is then calculated by the median of the GC-corrected fractions for the selected bins on chromosome Y. DEFRAG is available as an addition to WISECONDOR at https://github.com/rstraver/wisecondor.

BAYINDIR

Bayindir *et al.*[5] introduced a method that makes use of a robust estimate of the read count on X, Y and autosomal chromosomes. They divide the genome in 50 kb bins, and the read count fractions for every bin are determined. Then, they define two fractions:

$$BAYINDIR_a = \frac{med(Chr_{auto}) - med(Chr_X)^*}{med(Chr_{auto})} 2 \quad (2)$$

where $med(Chr_{auto})$ represents the median read count across the 50-kb bins of the autosomal chromosomes (and $med(Chr_X)$ likewise for bin on the X chromosome). BAYINDIR proposed a second method that only takes regions on the Y chromosome into account for which no read from female pregnancies map:

$$BAYINDIR_B = \frac{med(Chr_Y^{mask})}{med(Chr_{auto})} * 2 \quad (3)$$

where Chr_Y^{mask} excludes 21 50-kb regions on Y for which reads from female pregnancies map. A training set of 196 samples was used to train this method before applying it to our test samples.

SEQFF

The multivariate model of SeqFF[6] contains two regression models (elastic net and weighted rank selection criterion, WRSC), trained on an anonymized set of 25 312 male pregnancy

 © 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.

samples. Briefly, the genome is divided into 50-kb bins. Then, a weighted linear regression model that combines the read count fractions across all autosomal bins is learned to predict the read count for the Y chromosome (in the elastic net model), or the read counts for the separate bins on the Y chromosome taking their covariation into account (for the WRSC method). The fetal fraction is then the average of the predictions of the two models. Note, that although learned on Y, the authors have shown that the model also predicts fetal fractions for female pregnancies correctly. Here, we have used the pre-trained models from Reference 6 and used them to predict the fetal fraction in our samples.

SANEFALCON

SANEFALCON[7] determines the fetal fraction through the distribution of reads mapped around nucleosome positions on autosomal chromosomes. Therefore, this method is independent of the fetal gender. Briefly, a genome-wide nucleosome profile is generated by aligning all read count profiles with respect to detected nucleosome positions.

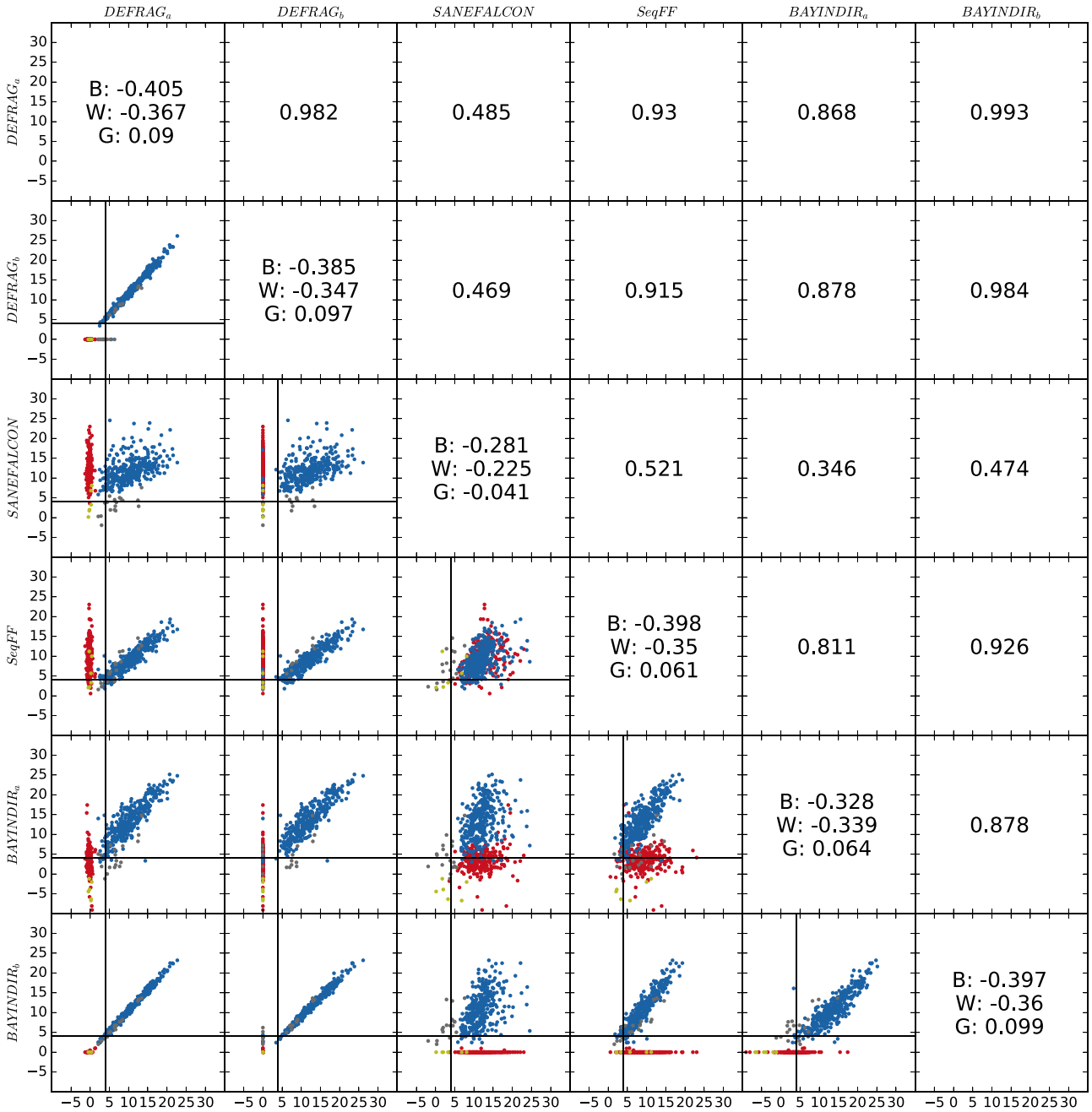

Figure 1 The lower left part of the matrix shows our comparison of the six different methods to predict fetal fraction from single read NGS data for 654 maternal blood plasma samples. Blue dots represent the male pregnancies, red dots the female pregnancies. Gray and green dots represent male and female pregnancies of a failed run, which contained degraded fetal DNA. On the diagonal of the matrix, the correlations to BMI (B), weight (W) and the gestational age (G) are shown, respectively. The upper right part of the matrix shows the correlation between the two methods (these data are also supplied separately as Supplementary Table 1)

© 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.

Nucleosome positions are detected by peaks in the genome-wide read count profile, as cfDNA are short fragments that are cut at linker DNA sites. The genome-wide nucleosome profile is then the averaged read count profile of the aligned profiles. Nucleosome profiles differ slightly between fetal enriched and maternal DNA due to differences in DNA degradation, i.e. fragment length. These changes correlate with the fetal fraction. SANEFALCON uses a linear regression from the nucleosome profile to predict the fetal fraction, with coefficients learned from a training set. Here, we have used the pre-trained model from the original publication.

RESULTS

High correlation between Y-based fetal fraction estimates

Figure 1 shows the correlations between the six different methods on the 654 test samples. High correlations were observed between methods that depend on the read count on chromosome Y, i.e. $DEFRAG_{a/b}$ and $BAYINDIR_{a/b}$.

Using unique male regions for Y chromosome improves fetal fraction estimate

$BAYINDIR_a$ predicts non-zero fetal fractions in female pregnancies (red samples), see for example the comparison with SeqFF. This is due to the reads mapping on the Y chromosome even for female pregnancies. A similar effect, although smaller, can be seen for the $DEFRAG_a$ method. When excluding those regions on the Y chromosome, as is done in $BAYINDIR_b$ and $DEFRAG_b$, we see that the predicted fetal fraction for female pregnancies is indeed zero. From this, we conclude that the methods that use only unique male regions on the Y chromosome to estimate the fetal fraction are to be preferred.

Fetal fraction for female pregnancies can be predicted, but less accurate

Two methods, SANEFALCON and SeqFF, can additionally detect fetal fraction for female pregnancies. We do observe a correlation between the two methods for the male pregnancies (blue samples, $\varrho = 0.52$) and the female pregnancies (red samples, $\varrho = 0.37$), suggesting that fetal fractions for female pregnancies can be estimated (see also Figure 1). Nevertheless, although both methods do correlate ($\varrho = 0.45$), their agreement is much less than the agreement between the Y-based methods ($0.87 < \varrho < 0.98$), indicating a lower accuracy for the fetal fraction for both SeqFF and SANEFALCON, with SeqFF performing slightly better than SANEFALCON.

Detection of technical failure

The green and gray samples in Figure 1 are from an NGS validation run in which an incidental human error occurred during sample preparation. As the wrong buffer was used, only large (maternal) fragments were selected, causing depletion of fetal DNA in the entire series. This was initially detected because several of the trisomy positive samples from this run were missed. Follow-up experiments showed that before this step an SRY signal could be determined for the male samples; after preparation, the signal was lost for 7 of the 18 male samples. From the Y-based methods, $DEFRAG_b$ performs best in detecting a (too) low fetal fraction, reporting such a (too) low fraction in 38.9% of the male samples. Interestingly, the comparable $BAYINDIR_b$ method is not as good in detecting these erroneous samples (27.8%). Remarkably, SANEFALCON is best in detecting the erroneous samples (61.1% for male samples and 66.7% for female samples), much better than SeqFF (33.3% for male samples and 50.0% for female samples). An explanation for this phenomenon lies in the SANEFALCON approach, which is based on the distribution of read-starts around nucleosome positions, where starts of longer reads (linked to maternal DNA) are further away from the nucleosomes. As a result, selecting for larger fragments has an immediate negative effect on the fetal fraction. The Y-based methods BAYINDIR and SeqFF performed less in detecting the erroneous samples, which might be explained by the fact that some of the male samples still contained sufficient Y-chromosomal DNA to conclude that a fetal fraction is present. Concluding, SANEFALCON can best be used to detect (too) low fetal fractions. Note that we have used a threshold of 4% as being proposed in literature to reduce the number of false negatives.[3] However, when inspecting Figure 1, the threshold for $DEFRAG_b$ could be lowered to 3% whereas for SANEFALCON a threshold of 6% might be preferred.

Body mass index, weight and gestational age

The correlation between the fetal fraction determined by the six methods and the weight (W), body mass index (B) and gestational age (G) are shown in the diagonal of Figure 1. Comparing these values with the results from Zhou *et al.*,[4] we observe a slightly stronger correlation between $DEFRAG_b$ and BMI than reported in Zhou *et al.*,[4] this might be caused by the use of different datasets. Regression of weight to $DEFRAG_b$ shows that for each extra kilogram, a decrease of 0.13% in fetal fraction is observed (0.42% decrease for each BMI point). Remarkably, no sample with a high BMI in the test set failed due to too low fetal fraction. Linear regression analysis on the gestational age further shows a positive correlation (+0.23% each week) with the fetal fraction.

DISCUSSION AND CONCLUSIONS

Our data confirm that it is important to determine the fetal fraction directly on the NGS data, as a separate assay may miss errors that are introduced during preparation for NGS. Furthermore, it is faster and more cost effective; all methods that were tested can easily be implemented into in-house pipelines and do not significantly increase the turnaround time. We found that the Y-based methods to estimate the fetal fraction are much more in agreement than the methods that are also capable to estimate the fetal fraction for female pregnancies. This agreement increased further when only male-specific regions on the Y-chromosome were considered; i.e. methods $DEFRAG_b$ and $BAYINDIR_b$. When considering the failed run, the $DEFRAG_b$ method was much better capable of identifying those samples that had a low fetal fraction. As reported by others, we found that BMI and gestational age do not correlate strongly with fetal fraction. Interestingly,

 © 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.

$DEFRAG_b$ shows the highest correlation, its correlation with BMI being even higher than previously reported correlations of BMI with fetal fraction. This strengthens our opinion that $DEFRAG_b$ gives the most accurate estimate of the fetal fraction, but its use is limited to male pregnancies. Both methods that also predict fetal fraction for female pregnancies agree on the fetal fractions for male and female pregnancies (similar correlations). However, they are less accurate than the Y-based methods (low correlation with the Y-based methods), with SeqFF performing slightly better than SANEFALCON. On the other hand, SANEFALCON performed best in detecting samples with very low fetal fractions from the failed experiment, even outperforming $DEFRAG_b$. This is extremely useful because in clinical practice we do not know whether there is a male or female pregnancy. Finally, we performed subgroup analysis to determine whether the performance of the different methods was dependent on the number of reads. As can be seen from Supplementary Figure 1, this was not the case. Taken together we propose the following for NIPT using a NGS single read sequencing strategy: (1) use SANEFALCON to check for low fetal fraction in all pregnancies. For samples that pass this check: (2) use $DEFRAG_b$ to estimate the fetal fraction for male pregnancies ($DEFRAG_b \geq 4\%$), and (3) use SeqFF to estimate the fetal fraction for female pregnancies ($DEFRAG_b < 4\%$). This proposal is based on our results on a relatively limited dataset of 654 samples and should be confirmed in future experiments.

WHAT'S ALREADY KNOWN ABOUT THIS TOPIC?

- It is important to determine fetal fraction during NIPT analysis, as a low fetal fraction may lead to false negative results.
- Several tools for the calculation of fetal fraction have been described.

WHAT DOES THIS STUDY ADD?

- A new tool for the calculation of fetal fraction, and a comparison of several of the existing tools.
- A clear recommendation on which tools should be used by laboratories performing NIPT analysis.
- New data on the influence of BMI and gestational age.

REFERENCES

1. Tabor A, Vestergaard CH, Lidegaard Ø. Fetal loss rate after chorionic villus sampling and amniocentesis: an 11-year national registry study. Ultrasound Obstet Gynecol 2009;34(1):19–24.
2. Tamminga S, van Maarle M, Henneman L, *et al.* Maternal plasma DNA and RNA sequencing for prenatal testing. Adv Clin Chem 2016;74:63–102.
3. Palomaki GE, Kloza EM, Lambert-Messerlian GM, *et al.* DNA sequencing of maternal plasma to detect down syndrome: an international clinical validation study. Genet Med 2011;13(11):913–20.
4. Zhou Y, Zhu Z, Gao Y, *et al.* Effects of maternal and fetal characteristics on cell-free fetal DNA fraction in maternal plasma. Reprod Sci 2015;22 (11):1429–35.
5. Bayindir B, Dehaspe L, Brison N, *et al.* Noninvasive prenatal testing using a novel analysis pipeline to screen for all autosomal fetal aneuploidies improves pregnancy management. Eur J Hum Genet 2015;23(10):1286–93.
6. Kim SK, Hannum G, Geis J, *et al.* Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. Prenat Diagn 2015;35(8):810–15.
7. Straver R, Oudejans CB, Sistermans EA, Reinders MJ. Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles. Prenat Diagn 2016;36(7):614–21.
8. Yu SC, Chan KC, Zheng YW, *et al.* Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. Proc Natl Acad Sci U S A 2014;111(23):8583–8.
9. Oepkes D, Page-Christiaens GC, Bax CJ, *et al.* Trial by Dutch laboratories for evaluation of non-invasive prenatal testing. Part I – clinical impact. Prenat Diagn Dec 2016;36(12):1083–90.
10. Straver R, Sistermans EA, Holstege H, *et al.* WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. Nucleic Acids Res 2014;42(5, e31).

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

© 2017 The Authors. *Prenatal Diagnosis* published by John Wiley & Sons, Ltd.