

Subject Section

# Combining count- and length-based z-scores leads to improved predictions in non-invasive prenatal testing

Jaroslav Budis<sup>1,2,3</sup>, Juraj Gazdarica<sup>2,4</sup>, Jan Radvanszky<sup>2,5</sup>, Gabor Szucs<sup>6</sup>, Marcel Kucharik<sup>7</sup>, Lucia Strieskova<sup>2,4</sup>, Iveta Gazdaricova<sup>4</sup>, Maria Harsanyova<sup>2,4</sup>, Frantisek Duris<sup>2,3,\*</sup>, Gabriel Minarik<sup>7</sup>, Martina Sekelska<sup>7</sup>, Balint Nagy<sup>8</sup>, Jan Turna<sup>3,4,9</sup>, and Tomas Szemes<sup>2,4,9</sup>

<sup>1</sup>Department of Computer Science, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia

<sup>2</sup>Geneton s.r.o., Bratislava, Slovakia

<sup>3</sup>Slovak Centre of Scientific and Technical Information, Bratislava, Slovakia

<sup>4</sup>Department of Molecular Biology, Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia

<sup>5</sup>Institute of Clinical and Translational Research, Biomedical Research Center, Slovak Academy of Sciences Bratislava, Slovakia

<sup>6</sup>Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia

<sup>7</sup>Medirex a.s., Bratislava, Slovakia

<sup>8</sup>Department of Human Genetics, University of Debrecen, Debrecen, Hungary

<sup>9</sup>Comenius University Science Park, Bratislava, Slovakia

\*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

## Abstract

**Motivation:** Non-invasive prenatal testing or NIPT is currently among the top researched topic in obstetric care. While the performance of the current state-of-the-art NIPT solutions achieve high sensitivity and specificity, they still struggle with a considerable number of samples that cannot be concluded with certainty. Such uninformative results are often subject to repeated blood sampling and re-analysis, usually after two weeks, and this period may cause a stress to the future mothers as well as increase the overall cost of the test.

**Results:** We propose a supplementary method to traditional z-scores to reduce the number of such uninformative calls. The method is based on a novel analysis of the length profile of circulating cell free DNA which compares the change in such profiles when random-based and length-based elimination of some fragments is performed. The proposed method is not as accurate as the standard z-score; however, our results suggest that combination of these two independent methods correctly resolves a substantial portion of healthy samples with an uninformative result. Additionally, we discuss how the proposed method can be used to identify maternal aberrations, thus reducing the risk of false positive and false negative calls.

**Availability:** The open-source code of the proposed methods, together with test data, is freely available for non-commercial users at github web page <https://github.com/jbudis/lambda>.

**Contact:** [fduris@dcs.fmph.uniba.sk](mailto:fduris@dcs.fmph.uniba.sk).

**Supplementary information:** Supplementary materials are available at *Bioinformatics* online.

## 1 Introduction

Prenatal screening and diagnostics are important parts of obstetric care. Current methods of prenatal testing still involve most commonly invasive sampling of fetal material using procedures such as amniocentesis and chorionic villus sampling which are associated with a small but real risk of miscarriage 0.5 – 1% (Mujezinovic and Alfirevic, 2007). To prevent the risk of abortion associated with invasive sampling procedures, non-invasive prenatal testing (NIPT) based on fetal DNA analysis from maternal circulation has been developed. In 1997, the discovery of fetal cell-free DNA (cfDNA) in maternal plasma and serum revolutionized the area of non-invasive prenatal diagnostics, and opened up new options in the field of obstetric research (Lo *et al.*, 1997). The fetal cfDNA is of placental origin (Bischoff *et al.*, 2005), and it can be reliably detected from fifth week of gestation (Lo *et al.*, 1998). On average, the fetal cfDNA contributes about 10% of all cfDNA fragments circulating in woman's blood when sampling is carried out between 10 and 20 gestational weeks, although the variance is quite large (Fiorentino *et al.*, 2016). The advance of massively parallel sequencing technologies together with the rapid development of bioinformatic algorithms and tools ushered in a new era of non-invasive prenatal identification of common fetal aneuploidies, now commonly known as NIPT (Chiu *et al.*, 2008; Fan *et al.*, 2008; Chiu *et al.*, 2011; Sehnert *et al.*, 2011; Bianchi *et al.*, 2012; Straver *et al.*, 2013; Yu *et al.*, 2014; Tynan *et al.*, 2016).

While the performance of the current state-of-the-art NIPT solutions achieve high sensitivity and specificity (Bianchi *et al.*, 2014; Koumbaris *et al.*, 2016), they still struggle with a considerable number of samples that cannot be concluded with certainty. The great source of such uninformative samples is in the nature of the statistical testing. Considering the standard cut-off threshold 2.5 in traditional z-score methods for reliable conclusion of healthy samples (Bianchi *et al.*, 2014), and testing normally distributed ratios measured for the common aneuploidy chromosomes, the chance that a healthy sample would achieve z-score greater than this is around 1.86%, and it is even higher when testing for other aberrations such as monosomy, gonosomal or sub-chromosomal aberrations.

Other problems are represented by maternal DNA aberrations such as maternal mosaicism (Wang *et al.*, 2014; Grati *et al.*, 2014), unidentified maternal tumours (Osborne *et al.*, 2013; Amant *et al.*, 2015), or copy number variations (Snyder *et al.*, 2015; Zhou *et al.*, 2017). Hypothetically, a duplication of even a small part of maternal chromosome, which may not be detrimental for the mother, may result in a false positive call for fetal aneuploidy. This is because such duplication effectively increases the size of that chromosome, and, because maternal cfDNA is by far dominant (Lo *et al.*, 2010; Rava *et al.*, 2014; Shubina *et al.*, 2017), the signal from partial maternal duplication can be interpreted as full fetal trisomy. Similarly, an opposite effect can cause a false negative result.

There is a growing body of studies addressing this issue (Wang *et al.*, 2014, 2015a,b). A particularly interesting venue of research focuses on qualitative differences between fetal and maternal cfDNA fragments, namely their lengths. It was previously reported that fetal fragments are on average shorter than maternal and that this information can be useful. Particularly, Fan *et al.* (2008, 2010) and Lo *et al.* (2010) were among the first to point out that fetal cfDNA fragments are on average shorter than maternal. Minarik *et al.* (2015) showed that *in silico* and *in vitro* cfDNA fragment size selection each leads to increase of fetal cfDNA fragments portion in the sample. Yu *et al.* (2014) defined a size ratio (i.e., ratio between abundances of fragments of certain lengths) of which they showed has high correlation with fetal fraction. Additionally, they defined size-based z-score as an alternative to the more common chromosome-count-based z-score. Zhang *et al.* (2017) proposed a modification of count-based z-score

(Chiu *et al.*, 2008) by multiplying the fractional genomic representation of the trisomic chromosome (i.e., 13<sup>th</sup>, 18<sup>th</sup> or 21<sup>st</sup>) by the ratio between cfDNA fragments shorter than 100, 130, 150 or 166 base pairs mapped to trisomic chromosome and reference chromosomes (all autosomes except 13<sup>th</sup>, 18<sup>th</sup> and 21<sup>st</sup>). Sun *et al.* (2017) proposed a method called COFFEE which does not require a set of reference samples to distinguish between T21 trisomic and euploid samples. Briefly, the method splits cfDNA fragments into two categories: shorter than 150bp and longer than 170bp. The chromosome 21 is also split into 200-kb bins, and a proportion of reads in short and long category for each bin is calculated. The status of the sample is determined based on the p-value of paired Mann-Whitney rank-sum test between cfDNA proportions in short and long category from the trisomic chromosome. Cirigliano *et al.* (2017) proposed a supposedly very robust method based on the likelihood ratios reflecting the probability for a sample to be affected versus the probability not to be affected given the sample's counting data (NCV score, Sehnert *et al.*, 2011) and fetal fraction. Rather unusually, no additional details were disclosed. However, the method seems to be somewhat reminiscent of Jiang *et al.* (2012) and Tynan *et al.* (2016). Shubina *et al.* (2017) were able to identify, post test, which of the false positive trisomy X samples were due to maternal mosaicism. Briefly, they observed that in case of true fetal aneuploidy, the fetal fraction calculated from X chromosome increases when the long reads are filtered out. On the other hand, for maternal mosaicism the filtering has almost no effect on this fetal fraction.

For the sake of completeness, we note that there are other qualitative information in each cfDNA fragment, for example its sequence content. Sung *et al.* (2015) showed that there is a difference in dispersion of short and long fragments across the human genome. Since fetal fragments are on average shorter than maternal, this indirectly implies that there is difference in sequence content of fetal and maternal cfDNA (all analyses use only uniquely mapped fragments). However, it is difficult to train such a model because one cannot easily distinguish which fragments are fetal and which maternal. Sung *et al.* (2015) did not have this problem because they fed their model normalized read counts per 50-kb bins (i.e., they did not distinguish the source), so they did not use the sequence content directly. One way out of this problem is to use non-pregnant women as a source of maternal cfDNA content, and chromosome-Y-mapped cfDNA fragments from male pregnancies as a source of fetal content. However, we have doubts regarding the representativeness of the Y-based fetal source. For a comprehensive review of fetal DNA fraction estimation, see (Peng and Jiang, 2017).

Extending the above mentioned work, in this paper we present a novel method of combining count-based and length-based z-scores which improves on the similar state-of-the-art methods and also further boosts the elimination of uninformative results which have a potential to cause needless stress to the parents, requiring repeated blood samplings and analyses. The associated increase of expenses, in turn, lower viability of the NIPT product. Reducing such cases is, therefore, of high interest in the area of NIPT. Our method is also based on the length of cfDNA fragments. Specifically, we take into account how fragments of a particular length contribute to the z-score calculated from some chromosomes as in (Sehnert *et al.*, 2011). Our approach has some similarities with previously published approaches, such as the fragments size distribution acquired from the pair-end sequencing from Yu *et al.* (2014), length-based modification of count-based z-scores from Minarik *et al.* (2015), combination of count-based and length-based information from Zhang *et al.* (2017) although the proposed combination is performed in a robust statistical manner, contribution of short and long cfDNA fragments to overall number of mapped fragments to the chromosome of interest from Sun *et al.* (2017).

Even if some parts of the proposed methods are not as accurate as the standard z-score, our results suggest that combination of these two independent methods correctly resolves a substantial portion of healthy

samples with an uninformative result. Additionally, we discuss application of the novel method for distinguishing between fetal (e.g., aneuploidy) and maternal signal (e.g., copy number variation).

## 2 Material and methods

### 2.1 Sample acquisition

We have collected altogether 2,621 samples with singleton pregnancy, of which 2 569 were negative for trisomy of chromosomes 13, 18 and 21, while 5 were confirmed as T13, 6 were confirmed as T18, and 39 were confirmed as T21. Positive results were in each case confirmed by amniocentesis. Negative samples were, however, not confirmed by any additional gold standard method. Data analyses reported here, were, on the other hand performed only on samples originally analysed with a sufficient time interval to know, from a clinician feedback following the delivery, whether any false negative results occurred.

One negative sample (analysed twice) was falsely reported as T18 (discovered by post-test amniocentesis). The samples were predominantly of Slovak and Czech origin. All women participating in this study gave informed written consent consistent with the Helsinki declaration.

### 2.2 Sample preparation and sequencing

Blood from pregnant women was collected into EDTA tubes and kept at 4°C temperature until plasma separation. Blood plasma was separated within 36 hours after collection and stored at -20°C until DNA isolation. DNA was isolated using Qiagen DNA Blood Mini kit. Standard fragment libraries for massively parallel sequencing were prepared from isolated DNA using an Illumina TruSeq Nano kit and a modified protocol described previously (Minarik *et al.*, 2015). Briefly, to decrease laboratory costs, we used reduced volumes of reagents what was compensated by 9 cycles of PCR instead of 8 as per protocol. Physical size selection of cfDNA fragments was performed using specific volumes of magnetic beads in order to enrich fetal fraction. Illumina NextSeq 500/550 High Output Kit v2 (75 cycles) was used for massively parallel sequencing of prepared libraries using pair-end sequencing with read length of 2x35bp on an Illumina NextSeq 500 platform.

### 2.3 Mapping and read count correction

Sequencing reads were aligned to the human reference genome (hg19) using Bowtie 2 algorithm (Langmead *et al.*, 2009). The first stage of data processing was carried out as previously described (Minarik *et al.*, 2015). NextSeq-produced fastq files (two per sample; R1 and R2) were directly mapped using the Bowtie 2 algorithm with very-sensitive option. Unless stated otherwise, only randomly chosen 5 million of alignments for each sample were considered, thus reducing the between-sample variability induced by sequencing. Note that all our samples had at least 5m reads. Reads with mapping quality of 40 or higher were retained for further data processing. Next, for each sample the unique reads were processed to eliminate the GC bias according to (Liao *et al.*, 2014) with the exclusion of intrarun normalization. Briefly, for each sample the number of unique reads from each 20kbp bin on each chromosome was counted. With empty bins filtered out, the locally weighted scatterplot smoothing (LOESS) regression was used to predict the expected read count for each bin based on its GC content. The LOESS-corrected read count for a particular bin was then calculated as  $RC_{cor} = RC - |RC_{loess} - RC_{avg}|$ , where  $RC_{avg}$  is the global average of read counts through all bins,  $RC_{loess}$  is the fitted read count of that bin, and  $RC$  is its observed read count.

Furthermore, variability of human genome in population also attributes to the mapping bias, mainly in regions with common structural differences.

Some methods therefore compare bin counts across a control set of sequenced samples to reduce weight of highly variable bins (Johansson *et al.*, 2017). A promising approach is based on principal component analysis (Price *et al.*, 2006; Zhao *et al.*, 2015). At first, bin counts are transformed into a principal space. The first component represents the highest variability across individuals in the control set. To normalize the sample, bin counts corresponding to predefined number of top components are removed to reduce common noise in euploid samples.

### 2.4 Reference z-score calculation

The reference z-scores of samples were calculated as normalized chromosome values (NCV) according to Sehnert *et al.* (2011). Given our training set, the optimal reference chromosomes were determined to be 1, 4, 8, 10, 19 and 20 for trisomy 21, 4, 7, 8, 9, 10 and 16 for trisomy 18, and 3, 4 and 7 for trisomy 13. Similarly to (Bianchi *et al.*, 2012), samples scoring 4 and higher were considered trisomic, while samples scoring 2.5 or lower were considered euploid. The range (2.5, 4) was considered uninformative. We will refer to these NCV values as reference z-scores or  $Z_{NCV}$ .

### 2.5 Length score calculation

We defined three novel statistics based on the fragment lengths, each building on the previous one. The basis for our first novel statistic, termed  $\lambda$ -score, were read counts for the chromosome of interest (e.g., the usual 13<sup>th</sup>, 18<sup>th</sup> or 21<sup>st</sup> chromosome). However, in contrast with the traditional z-scores, we did not compare the read count with the expected normal value estimated from a set of euploid samples. Instead, we continuously eliminated fragments of certain length and compared the observed counts with the expected counts, if the same amount of fragments was eliminated randomly.

More particularly, if  $i$  marks the chromosome of interest and  $chr_i$  the number of reads mapped to the  $i^{th}$  chromosome of the tested sample, then let  $chr_i(l)$  be the number of reads mapped to the  $i^{th}$  chromosome that are of length at most  $l$ . Furthermore, let the total number of reads of the sample be  $n$ , and let  $n_l$  be the total number of reads that are of length at most  $l$ . The number of reads mapped to the  $i^{th}$  chromosome (of any length) after uniform random elimination follows a binomial distribution with parameters  $(n_l, p)$ ,  $p = chr_i/n$ . The ratio  $chr_i/n$  gives the proportion of reads originally mapped to the  $i^{th}$  chromosome, and the number  $n_l$  gives the number of reads to be drawn. Thus, the expected number of reads mapped to the  $i^{th}$  chromosome after the uniform random draw and its variance is equal to  $e_l = n_l p$  and  $v_l = n_l p(1 - p)$  as in binomial distribution, respectively.

For each sample and chromosome, we defined a series of  $\lambda$ -scores as

$$\lambda_i(l) = \frac{chr_i(l) - n_l p}{\sqrt{n_l p(1 - p)}}, \quad (1)$$

where  $l$  ranges from 50 to 200 and  $i$  indicates a chromosome. Next, we defined a second novel statistic termed  $FL$ -score as

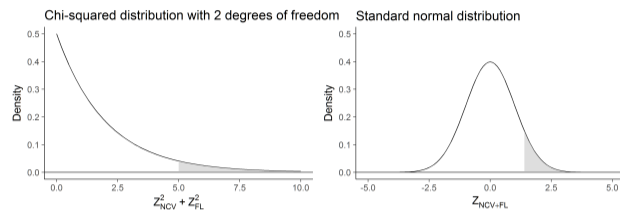
$$FL_i = \max_{125 \leq l \leq 145} \lambda_i(l). \quad (2)$$

The bounds 125 and 145 were determined empirically. Finally, the  $FL$ -scores were normalized into z-scores which approximately follow standard normal distribution for euploid samples (Supplementary figure S1). Then, for any sample, its normalized  $FL$ -score value  $Z_{FL}$  was used as an alternative method for the prediction of aneuploidy.

### 2.6 Combining the scores

In our analysis, we found no correlation between the reference ( $Z_{NCV}$ ) and length-based ( $Z_{FL}$ ) scores measured in our training set of euploid

samples (Pearson  $R = 0.017$ ,  $p = 0.129$ ). Thus, we considered them as two independent random variables, each from a standard normal distribution. The sum of their squares follows a chi-squared distribution with 2 degrees of freedom, and a survival function of the chi-squared distribution was used to associate this sum with the probability. Finally, this probability was converted back to a standard score  $Z_{NCV+FL}$  through quantile function for easy comparison with other methods (Figure 1). Note that the calculations were performed in log-space to overcome underflow issues.



**Fig. 1.** The grey area of the left plot represents the probability of observing  $Z_{NCV}^2 + Z_{FL}^2 \geq 5$ , which in this case is equal to 0.082. On the other hand, this probability is associated with z-score 1.391 in standard normal distribution in a way that the probability of observing z-score  $\geq 1.391$  is the same 0.082. Thus, the shaded areas are equal.

On the other hand, there was a significant correlation between  $Z_{SZ}$  (defined below in section 3.2) and  $Z_{FL}$  (Pearson  $R = 0.55$ ,  $p < 0.001$ ). In this case, we combined them according to (Owen and Chmielewski, 1985). Specifically, we first performed a principal component analysis of the pairs  $(Z_{SZ}, Z_{FL})$ . Then, we recalculated the z-scores along the newly found eigenvectors  $e_1, e_2$  using the respective eigenvalues as new variances. The resulting scores were two independent standard normals, and we proceeded to combine them as in the first paragraph of this section. The resulting z-score was marked as  $Z_{SZ+FL}$  (Figure 2).

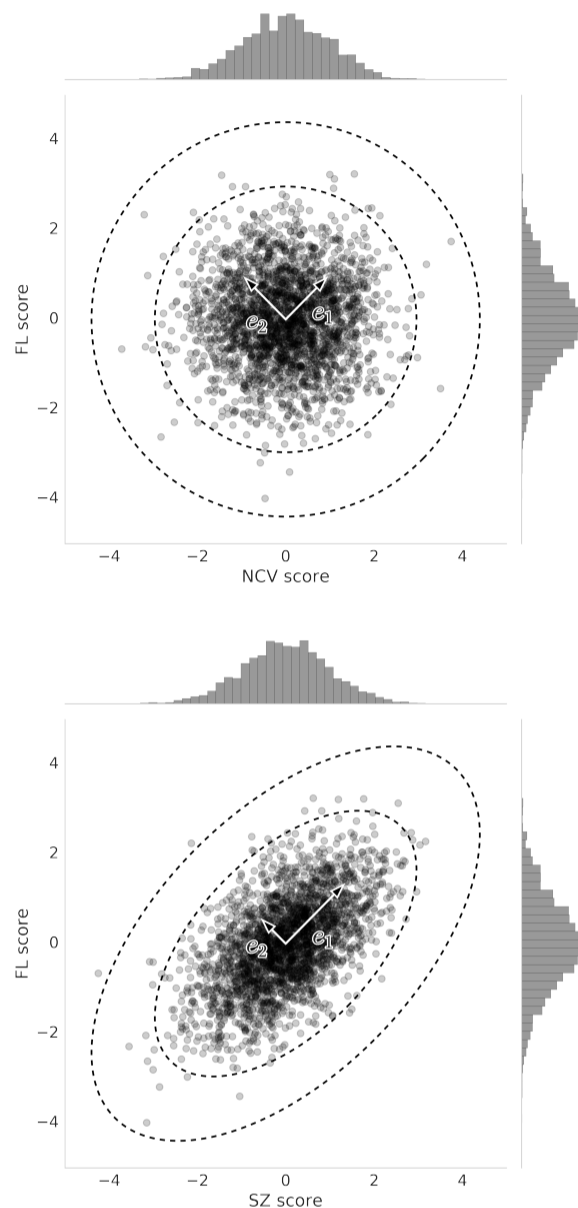
## 2.7 Statistical analysis

The significance of our findings was evaluated using statistical tests implemented in Python *scipy* package (Jones et al., 2014). The linear dependency of the two scores in negative samples was calculated with Pearson correlation. Since scores of aberrant samples were not normally distributed, Wilcoxon signed-rank test was used to estimate statistical significance of improvement between the reference and proposed methods.

## 3 Results

### 3.1 Lambda-score profiles

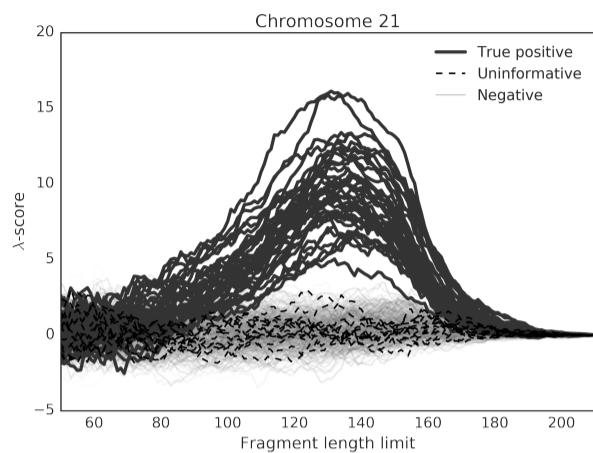
First, we calculated series of  $\lambda$ -scores for euploid and trisomic samples for chromosomes 13, 18 and 21, and length range from 50bp to 220bp. We observed that trisomic samples behaved differently than euploid samples (Figure 3). This difference can be explained by fetal fragments being shorter than maternal. The fetal and maternal cfDNA differ in properties such as length distribution and source chromosomes (Fan et al., 2010; Yu et al., 2014). Originally, they are mixed in some ratio, usually termed fetal fraction, resulting in the observed properties of the mixture (read length distribution and chromosome mapping ratios). Uniform random elimination of fragments preserves these properties of the mixture on account of being uniform, both in euploid and trisomic cases. On the other hand, elimination by length as described in section 2.5 eliminates more maternal than fetal fragments on account of the latter being shorter than the former. Thus, the original properties of the mixture are not preserved, if fetal and maternal properties are not the same, which is the case for trisomic samples with respect to chromosome mapping ratios.



**Fig. 2.** Combination of novel FL score with the reference score NCV (top) and length-reduced reference score SZ (bottom). Displayed are negative samples for chromosome 21. The ellipses represent combined z-scores 2.5 and 4, respectively. On the top plot, the scores were combined through  $\chi^2$  distribution with 2 degrees of freedom. On the bottom plot, first, principal component analysis was used to calculate axes  $(e_1, e_2)$  of variation along which the two principal components were independent. Subsequently,  $\chi^2$  distribution with 2 degrees of freedom was used to combine the principal components.

At first, the  $\lambda$ -scores of trisomic samples gradually increase with the elimination of longer fragments. This positive effect is however balanced by the negative effect of lower number of remaining fragments, and so, after while, the  $\lambda$ -scores decline to the values expected by a random draw. We observed the highest deviation of aberrant samples using thresholds for fragment lengths between 125bp and 145bp. We therefore measured the maximal value in this range which we termed FL score of the sample.

Note that in contrast with the reference z-score  $Z_{NCV}$ , the calculation of  $\lambda$ -scores is not based on comparison of proportions of fragments with healthy population. An excessive number of fragments from a chromosome



**Fig. 3.** Deviation of the T21 positive samples from the euploid samples in terms of  $\lambda$ -score. Elimination of fragments with length above a threshold (Fragment length limit) leads to divergence of chromosomal counts from the expectations. In the healthy samples (only 500 samples shown for clarity), the difference can be explained by fluctuations in random draw. On the other hand, aberrant samples show profile that markedly diverge from expected values.

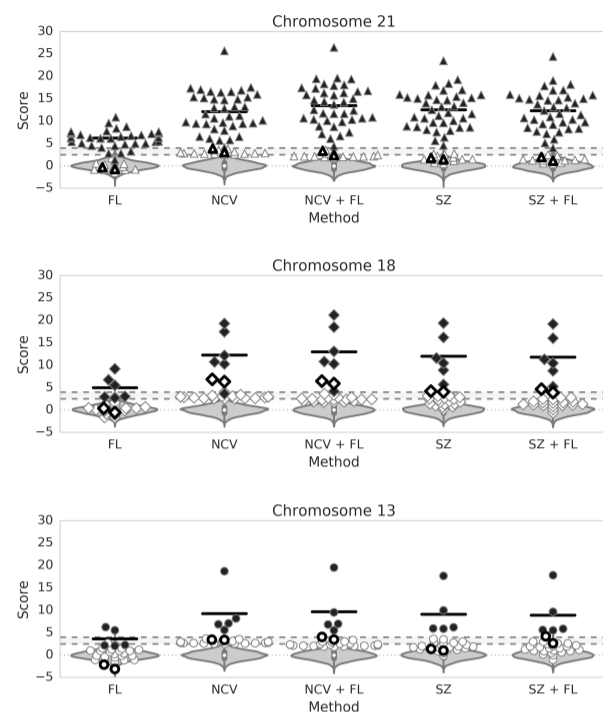
would therefore not result in the positive call, if these fragments do not have fetal length distribution. Thus, it is possible to reveal false positive calls caused by maternal aberrations. Similar concepts have been already utilized in distinction between maternal and fetal gonosomal aberrations (Shubina *et al.*, 2017), leading to the reduction of false positive results of the monosomy X0 predictions (section 3.5).

### 3.2 *In silico* size selection

Next, we examined the effect of the length-based fragment filtering on the reference z-scores. Particularly, we first calculated the reference z-scores  $Z_{NCV}$  (section 2.4), and then we removed the fragments longer than 150bp. However, this considerably changed the read count for our samples ( $1.5 \pm 0.3$  million), so we concluded that the trained mean and standard deviation (used in z-score calculation) for the original data may not be suitable for the length-reduced data. Therefore, we applied this length-reduction to our training samples as well, determined the new mean and standard deviation (we kept the reference chromosomes the same), and only with these new values we calculated z-scores of the length-reduced test samples. Comparing the original and length-reduced z-scores (Figure 4), we observed only small and statistically insignificant increase in z-scores of trisomic samples (average multiplicative increase  $1.05\times$ ; Wilcoxon  $Z = 516$ ,  $p = 0.241$ ). For future reference, we termed these z-scores  $Z_{SZ}$ .

On the other hand, when all aligned reads were considered (recall that up until now, all samples were restricted to the first 5 million raw alignments, see section 2.3), and the same procedure was applied again, the increase in z-scores of trisomic samples became statistically significant (average multiplicative increase  $1.13\times$ ; Wilcoxon  $Z = 136$ ,  $p < 0.001$ ). As before, we recalculated the mean and standard deviation used in the z-score calculation because of the changed read count per sample ( $8.8 \pm 4.6$  million, min = 5m, max = 46.3m), while the reference chromosomes were again kept the same (section 2.4).

This finding indicates that the length-based fragment selection is beneficial only for samples with more than 5 million aligned fragments, at least for the reference z-scores. This is in accord with our previous findings that *in silico* size-based filtering of fragments (only reads up to 155bp were retained) did not lead to statistically significant increase in trisomic



**Fig. 4.** Comparison of z-score calculation methods on chromosome 21 (upper), chromosome 18 (middle) and chromosome 13 (lower). Reference method based on chromosomal fragment counts (NCV) performs better than proposed FL method, albeit their combination (NCV + FL) increased z-scores of trisomic samples (black marks). Size selection (SZ) further improves z-scores of trisomic samples. The combination with FL method (SZ + FL) most markedly reduces z-scores of samples evaluated as uninformative using traditional NCV score alone (white marks). Empty shapes represent results of the two analyses of the IVF sample discussed in section 3.5. See Supplementary figures S2, S3, S4 for larger plots which also include comparison with other methods as discussed in section 3.6. Scatterplot versions of some plots are shown on Supplementary figures S14 - S17.

z-scores using MiSeq runs having  $3.1 \pm 1.0$  million reads per sample, while statistically significant ( $p = 0.04$ ) increase was observed for samples sequenced on Ion Torrent PGM, of which samples had  $5.8 \pm 1.0$  million reads (Minarik *et al.*, 2015).

### 3.3 Combined scores as supplemental evaluation method

The performance of the novel statistic  $Z_{FL}$  by itself was observed to be weaker than that of the traditional  $Z_{NCV}$  (Figure 4). Particularly, there was a substantial decrease in z-scores for trisomic samples. On the other hand, their combination  $Z_{NCV+FL}$  resulted in statistically significantly higher trisomic z-scores than  $Z_{NCV}$  (average multiplicative increase  $1.11\times$ ; Wilcoxon  $Z = 17$ ,  $p < 0.001$ ). Surprisingly, the combined score  $Z_{SZ+FL}$  produced only statistically insignificant increase of trisomic z-scores (average multiplicative increase  $1.02\times$ ; Wilcoxon  $Z = 637$ ,  $p > 0.05$ ).

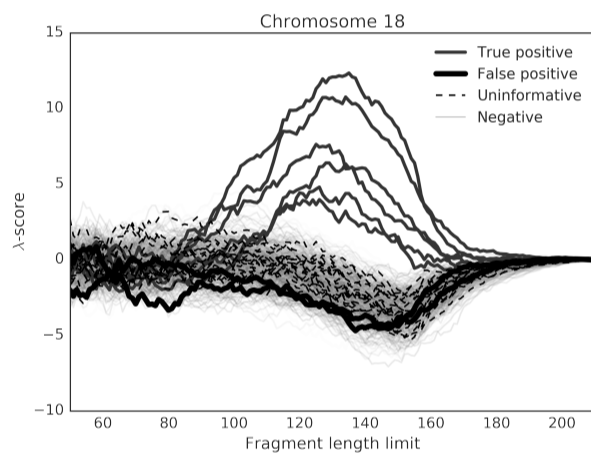
Furthermore, we observed a decrease of  $Z_{SZ+FL}$  scores of false positive and uninformative samples when compared with reference  $Z_{NCV}$ , even though other samples that were previously classified as negative replaced them in the uninformative range. Additionally,  $Z_{NCV+FL}$  and  $Z_{SZ+FL}$  resulted both in false positive results indicating that using these metrics alone may decrease accuracy of the testing.

### 3.4 Improved evaluation method

In light of our findings, we propose the following improvement of the NIPT evaluation process. First, a reference  $Z_{NCV}$  is calculated. If it is in the negative zone, then no further test is applied and the sample is closed as normal. If the  $Z_{NCV}$  is in the uninformative zone, then  $Z_{SZ+FL}$  is computed, and if this new score is in the negative zone, the sample is closed as normal. Our findings indicate that this process considerably decreases the number of uninformative results (Figure 4, the white marks in particular).

### 3.5 A false positive sample

We observed an atypical sample from IVF pregnancy (analysed twice) with high risk for trisomy 18 (Figure 5). Although both analyses were supported by solid  $Z_{NCV}$  scores 6.82 and 6.31, the predicted aberration was not confirmed by invasive follow-up test. On the other hand, the fragment length  $Z_{FL}$  score classified it as healthy sample (-0.57, 0.34). The combination of these scores  $Z_{NCV+FL}$  led to a slightly reduced scores (6.43, 5.88), and the inclusion of the *in silico* size selection further reduced their z-scores (3.64, 3.44), getting them into uninformative range.



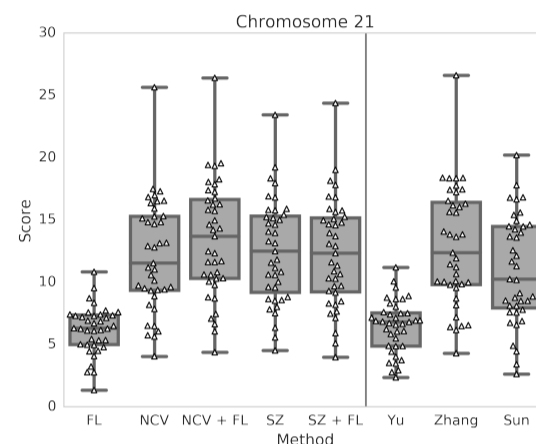
**Fig. 5.**  $\lambda$ -scores for chromosome 18. Interesting is the twice analysed false positive T18 sample discussed in section 3.5. The  $\lambda$ -score profile, similar to healthy samples, may indicate that the revealed aberration by  $Z_{NCV}$  (6.82 and 6.31) may be mistaken.

The most likely explanation is that this was a result of a maternal copy number variation on the 18<sup>th</sup> chromosome which is known to cause false positive results (Zhou et al., 2017). Unfortunately, we were not able to perform a follow-up of maternal genotype. If this was the case, it would indicate that the proposed method, particularly the  $\lambda$ -score profiles, can be used to distinguish between fetal (e.g., aneuploidy) and maternal (e.g., copy number variation) NIPT signals. Such method would be very valuable, but validation on a larger cohort of such samples is required.

Other possible and rather interesting hypothesis is that this was the case of a trisomic vanishing twin. In this case, we would expect that the cfDNA fragment length distribution originating from the trisomic vanishing twin would be similar to that of the trisomic living fetus, though we were not able to find publications supporting or opposing this expectation. However, the observation indicates that the fragment length distribution is similar to that of a mother. Thus, either the vanishing twin hypothesis or the assumption about the read length distribution is wrong. In case of the latter, such difference could then be utilized for reducing the false NIPT results due to vanishing twin effects.

### 3.6 Comparison with other methods

Using the trisomic dataset, we compared the proposed methods with three other methods which also use cfDNA fragment length distribution, particularly with methods due to Yu et al. (2014), Zhang et al. (2017) and Sun et al. (2017). The details of the computation of the three methods are in Supplementary material. Results for the T21 trisomic data set are shown on Figure 6. Similar results for chromosome 18 and 13 or shown on the Supplementary figures S5 and S6, respectively. Results including euploid as well as trisomic dataset are shown on Supplementary figures S2, S3 and S4 for chromosome 21, 18 and 13, respectively.



**Fig. 6.** Comparison of z-score calculation methods on chromosome 21 on the trisomic dataset. Yu stands for (Yu et al., 2014), Zhang stands for (Zhang et al., 2017) and Sun stands for (Sun et al., 2017).

We observed that  $Z_{Yu}$  behaves very similarly to  $Z_{FL}$  (Supplementary figures S7), although there was a statistically significant increase in trisomic z-scores for the  $Z_{Yu}$  method (average multiplicative increase 1.05 $\times$ ; Wilcoxon  $Z = 391$ ,  $p = 0.017$ ). On the other hand,  $Z_{Yu}$  performed much worse when compared with  $Z_{NCV+FL}$  (average multiplicative increase 2.13 $\times$ ; Wilcoxon  $Z = 0$ ,  $p < 0.001$ ) (Supplementary figure S8). Please note that this makes perfect sense because  $Z_{FL}$  and  $Z_{Yu}$  are both based solely on the fragment length information. On the other hand,  $Z_{NCV+FL}$  is based on chromosome-count information as well, and it greatly outperforms them both.

We also observed that  $Z_{Zhang}$  performed similarly to  $Z_{NCV+FL}$  (Supplementary figure S9),  $Z_{NCV+FL}$  although our method was marginally better (average multiplicative increase 1.05 $\times$ ; Wilcoxon  $Z = 94$ ,  $p < 0.001$ ). Again, this makes perfect sense. Both  $Z_{Zhang}$  and  $Z_{NCV+FL}$  are based on length and chromosome-count information, and they perform similarly.

COFFEE is difficult to compare as the  $-\log(p)$  values are not z-scores (especially visible for euploid samples on Supplementary figure S2). However, it has comparable values with  $Z_{NCV+FL}$  on T21 trisomic samples (Supplementary figure S10).

## 4 Conclusion

The reference z-score method ( $Z_{NCV}$ ) of aneuploidy prediction based on chromosomal proportions is already widely accepted, since it can quite well separate between trisomic and euploid samples. To avoid false predictions in routine diagnosis, NIPT tests have typically a range of scores

which are considered too risky for definite predictions. Such uninformative results are often subject to repeated blood sampling and re-analysis, usually after two weeks, and this period may cause a stress to the future mothers. In addition to that, high numbers of uninformative results also increase the overall cost of the whole procedure. The supplementary scores proposed in this paper may offer a useful way for reducing the number of uninformative samples in several ways.

In our data set, we classified 53 of 2,569 (2.06%) negative samples as uninformative using the reference  $Z_{NCV}$  method. The combination of  $Z_{NCV}$  with the  $Z_{FL}$  score,  $Z_{NCV+FL}$ , led to similar number 51 (1.98%), albeit only 15 of them were shared in both metrics. The addition of size selection to the combined method ( $Z_{SZ+FL}$ ) produced similar number of uninformative samples to 48 (1.87%) but only 4 of them were shared with the reference method. Thus, the combination of these approaches can be used to substantially reduce the number of uninformative results as proposed in section 3.4. In our case, 49 out of 53 uninformative samples would be closed as negative without affecting the prediction of true positive samples.

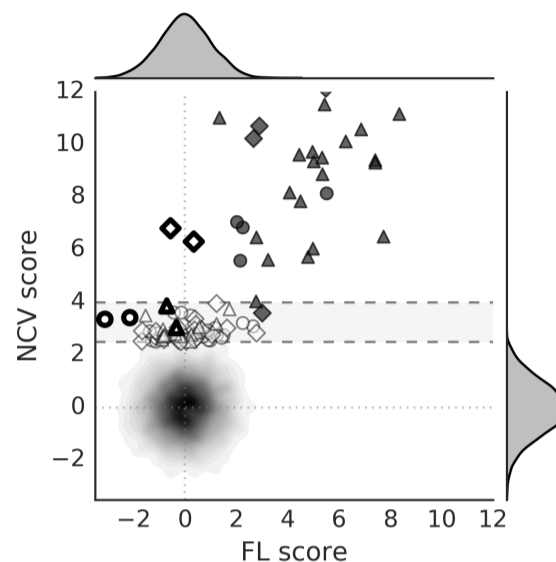
A proper verification of the proposed method should also include true positive samples with  $Z_{NCV}$  in the uninformative range which the proposed method would classify either as positive or uninformative (that is, it should not create false negatives). Unfortunately, we did not observe any true positive sample in the uninformative range. To address this issue, we subsampled our data to 2.5m reads (half of considered 5m earlier) what produced some true positive samples in the uninformative range for  $Z_{NCV}$  method. We observed that  $Z_{NCV+FL}$  method either correctly classified these samples as positive or kept them in the uninformative range (Supplementary figure S11, S12, S13). Note that we did not consider  $Z_{SZ+FL}$  method because this method (particularly  $Z_{SZ}$ ) requires at least 5m reads to work as discussed in section 3.2. While not conclusive, this is at least some kind of evidence that the proposed method does not create false negative calls out of uninformative results.

Another way to lower the number of uninformative samples is to classify uninformative samples with low  $Z_{FL}$  score as negative. Setting the threshold for  $Z_{FL}$  score to 0, 1, 2 led to correct elimination of 24 (45.28%), 41 (77.36%) and 50 (94.34%) out of 55 true negative samples with uninformative  $Z_{NCV}$  call. The thresholds, however, must be chosen with caution, since the lowest observed  $Z_{FL}$  score of positive sample was 1.34. Because this sample also had high chromosomal  $Z_{NCV}$  score 11.02, the false negative call with  $Z_{FL}$  threshold set to 2 could be avoided by preferring the reference method, albeit similarly low  $Z_{FL}$  scores may occur along with the low score of the reference method.

A typical NIPT analysis is suitable for prediction of monosomy as well. The monosomic samples are distinguished by negative z-scores below some predefined threshold, for example  $-3$  (Mazloom *et al.*, 2013). In contrast with the sign of chromosomal  $Z_{NCV}$  score, which indicates increase or decrease of DNA material from specific chromosome, the sign of the fragment length  $Z_{FL}$  score indicates maternal or fetal origin of the aberration. This way, the scores may be divided into four categories representing: 1) maternal duplication ( $Z_{FL} < 0$ ,  $Z_{NCV} > 0$ ), 2) fetal duplication ( $Z_{FL} > 0$ ,  $Z_{NCV} > 0$ ), 3) maternal deletion ( $Z_{FL} < 0$ ,  $Z_{NCV} < 0$ ), and 4) fetal deletion ( $Z_{FL} > 0$ ,  $Z_{NCV} < 0$ ), thus making the prediction more informative (Figure 7).

Also, we observed a statistically significant increase in trisomic z-scores for method  $Z_{NCV+FL}$ . Thus, this method can be considered as an improved version of  $Z_{NCV}$ .

Finally, we discuss the possibility of either our method being able to distinguish between fetal and maternal NIPT signals or trisomic vanishing twin having different cfDNA fragment length distribution than living trisomic fetus. To verify these hypotheses, more tests on a larger cohort of samples are required.



**Fig. 7.** Comparison of z-scores calculated chromosomal count (NCV) and fragment length (FL) methods. Positive (black) and negative (white) samples with z-scores higher than 2.5 (NCV method) are represented by triangles (chromosome 21), diamonds (chromosome 18) and circles (chromosome 13). Other negative samples are accumulated in form of a density plot in the intersection of major axes ( $x = 0$ ;  $y = 0$ ). Major axes divide predictions into four classes, while the fetal trisomies are located in top, right quartile ( $Z_{FL} > 0$ ,  $Z_{NCV} > 0$ ). Empty shapes represent results of the two analyses of the IVF sample discussed in section 3.5.

Elimination of long fragments may significantly improve prediction accuracy of trisomy testing. The number of sequenced DNA fragments must be however sufficient to balance the lower number of analysed fragments. Based on the patterns observed in z-score profiles, we designed a novel method for prediction that is independent to standard method based on chromosomal counts. We presented that combination of these two methods may conclude samples that cannot be safely classified using a single method. However, care should be taken when interpreting short cfDNA fragments as solely or predominantly fetal, because Jiang *et al.* (2015) showed that “populations of aberrantly short and long DNA molecules existed in the plasma of patients with hepatocellular carcinoma”. On the other hand, patients with carcinoma have other tell-tale signs such as large copy number aberrations that can help identify such case (Chan *et al.*, 2012; Dharajiya *et al.*, 2017).

## Funding

This contribution is the result of implementation of the project *REVO-GENE – Research centre for molecular genetics* (ITMS 26240220067) supported by the Research & Developmental Operational Programme funded by the European Regional Development Fund.

## References

- Amant, F., Verheecke, M., Wlodarska, I., Dehaspe, L., Brady, P., Brison, N., Van Den Bogaert, K., Dierickx, D., Vandecaveye, V., Tousseyn, T., *et al.* (2015). Presymptomatic identification of cancers in pregnant women during noninvasive prenatal testing. *JAMA oncology*, **1**(6), 814–819.
- Bianchi, D. W., Platt, L. D., Goldberg, J. D., Abuhamad, A. Z., Sehnert, A. J., and Rava, R. P. (2012). Genome-wide fetal aneuploidy detection by maternal plasma dna sequencing. *Obstetrics & Gynecology*, **119**(5), 890–901.

- Bianchi, D. W., Parker, R. L., Wentworth, J., Madankumar, R., Saffer, C., Das, A. F., Craig, J. A., Chudova, D. I., Devers, P. L., Jones, K. W., et al. (2014). Dna sequencing versus standard prenatal aneuploidy screening. *New England journal of medicine*, **370**(9), 799–808.
- Bischoff, F. Z., Lewis, D. E., and Simpson, J. L. (2005). Cell-free fetal dna in maternal blood: kinetics, source and structure. *Human reproduction update*, **11**(1), 59–67.
- Chan, K. A., Jiang, P., Zheng, Y. W., Liao, G. J., Sun, H., Wong, J., Siu, S. S. N., Chan, W. C., Chan, S. L., Chan, A. T., et al. (2012). Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clinical chemistry*, pages clinchem–2012.
- Chiu, R. W., Chan, K. A., Gao, Y., Lau, V. Y., Zheng, W., Leung, T. Y., Foo, C. H., Xie, B., Tsui, N. B., Lun, F. M., et al. (2008). Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of dna in maternal plasma. *Proceedings of the National Academy of Sciences*, **105**(51), 20458–20463.
- Chiu, R. W., Akolekar, R., Zheng, Y. W., Leung, T. Y., Sun, H., Chan, K. A., Lun, F. M., Go, A. T., Lau, E. T., To, W. W., et al. (2011). Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma dna sequencing: large scale validity study. *Bmj*, **342**, c7401.
- Cirigliano, V., Ordoñez, E., Rueda, L., Syngelaki, A., and Nicolaidis, K. (2017). Performance of the neobona test: a new paired-end massively parallel shotgun sequencing approach for cell-free dna-based aneuploidy screening. *Ultrasound in Obstetrics & Gynecology*, **49**(4), 460–464.
- Dharajiya, N. G., Grosu, D. S., Farkas, D. H., McCullough, R. M., Almasri, E., Sun, Y., Kim, S. K., Jensen, T. J., Saldivar, J.-S., Topol, E. J., et al. (2017). Incidental detection of maternal neoplasia in noninvasive prenatal testing. *Clinical chemistry*, pages clinchem–2017.
- Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L., and Quake, S. R. (2008). Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing dna from maternal blood. *Proceedings of the National Academy of Sciences*, **105**(42), 16266–16271.
- Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L., and Quake, S. R. (2010). Analysis of the size distributions of fetal and maternal cell-free dna by paired-end sequencing. *Clinical chemistry*, **56**(8), 1279–1286.
- Fiorentino, F., Bono, S., Pizzuti, F., Mariano, M., Polverari, A., Duca, S., Sessa, M., Baldi, M., Diano, L., and Spinella, F. (2016). The importance of determining the limit of detection of non-invasive prenatal testing methods. *Prenatal diagnosis*, **36**(4), 304–311.
- Grati, F. R., Malvestiti, F., Ferreira, J. C., Bajaj, K., Gaetani, E., Agrati, C., Grimi, B., Dulcetti, F., Ruggeri, A. M., De Toffol, S., et al. (2014). Fetoplacental mosaicism: potential implications for false-positive and false-negative noninvasive prenatal screening results. *Genetics in Medicine*, **16**(8), 620–624.
- Jiang, P., Ren, J., Chen, F., Zhou, Y., Xie, J., Dan, S., Su, Y., Xie, J., Yin, B., Su, W., et al. (2012). Noninvasive fetal trisomy (nifty) test: an advanced non-invasive prenatal diagnosis methodology for fetal autosomal and sex chromosomal aneuploidies. *BMC medical genomics*, **5**(1), 57.
- Jiang, P., Chan, C. W., Chan, K. A., Cheng, S. H., Wong, J., Wong, V. W.-S., Wong, G. L., Chan, S. L., Mok, T. S., Chan, H. L., et al. (2015). Lengthening and shortening of plasma dna in hepatocellular carcinoma patients. *Proceedings of the National Academy of Sciences*, **112**(11), E1317–E1325.
- Johansson, L. F., Boer, E., Weerd, H., Dijk, F., Elferink, M., Schuring-Blom, G., Suijkerbuijk, R. F., Sinke, R. J., Meerman, G., Sijmons, R. H., et al. (2017). Novel algorithms for improved sensitivity in non-invasive prenatal testing. *Scientific Reports*, **7**(1), 1838.
- Jones, E., Oliphant, T., and Peterson, P. (2014). {SciPy}: open source scientific tools for {Python}.
- Koumbaris, G., Kypri, E., Tsangaras, K., Achilleos, A., Mina, P., Neofytou, M., Velissariou, V., Christopoulou, G., Kallikas, I., González-Liñán, A., et al. (2016). Cell-free dna analysis of targeted genomic regions in maternal plasma for non-invasive prenatal testing of trisomy 21, trisomy 18, trisomy 13, and fetal sex. *Clinical chemistry*, **62**(6), 848–855.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short dna sequences to the human genome. *Genome biology*, **10**(3), R25.
- Lo, Y. D., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W., and Wainscoat, J. S. (1997). Presence of fetal dna in maternal plasma and serum. *The Lancet*, **350**(9076), 485–487.
- Lo, Y. D., Tein, M. S., Lau, T. K., Haines, C. J., Leung, T. N., Poon, P. M., Wainscoat, J. S., Johnson, P. J., Chang, A. M., and Hjelm, N. M. (1998). Quantitative analysis of fetal dna in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *The American Journal of Human Genetics*, **62**(4), 768–775.
- Lo, Y. D., Chan, K. A., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M., Zheng, Y. W., Leung, T. Y., Lau, T. K., Cantor, C. R., et al. (2010). Maternal plasma dna sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Science translational medicine*, **2**(61), 61ra91–61ra91.
- Mazloom, A. R., Džakula, Ž., Oeth, P., Wang, H., Jensen, T., Tynan, J., McCullough, R., Saldivar, J.-S., Ehrich, M., Boom, D., et al. (2013). Noninvasive prenatal detection of sex chromosomal aneuploidies by sequencing circulating cell-free dna from maternal plasma. *Prenatal diagnosis*, **33**(6), 591–597.
- Minarik, G., Repiska, G., Hyblova, M., Nagyova, E., Soltys, K., Budis, J., Duris, F., Sysak, R., Bujalkova, M. G., Vlkova-Izrael, B., et al. (2015). Utilization of benchtop next generation sequencing platforms ion torrent pgm and miseq in noninvasive prenatal testing for chromosome 21 trisomy and testing of impact of in silico and physical size selection on its analytical performance. *PLoS one*, **10**(12), e0144811.
- Mujezinovic, F. and Alfirevic, Z. (2007). Procedure-related complications of amniocentesis and chorionic villous sampling: a systematic review. *Obstetrics & Gynecology*, **110**(3), 687–694.
- Osborne, C. M., Hardisty, E., Devers, P., Kaiser-Rogers, K., Hayden, M. A., Goodnight, W., and Vora, N. L. (2013). Discordant noninvasive prenatal testing results in a patient subsequently diagnosed with metastatic disease. *Prenatal diagnosis*, **33**(6), 609–611.
- Owen, J. G. and Chmielewski, M. A. (1985). On canonical variates analysis and the construction of confidence ellipses in systematic studies. *Systematic Zoology*, **34**(3), 366–374.
- Peng, X. L. and Jiang, P. (2017). Bioinformatics approaches for fetal dna fraction estimation in noninvasive prenatal testing. *International journal of molecular sciences*, **18**(2), 453.
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, **38**(8), 904.
- Rava, R. P., Srinivasan, A., Schnert, A. J., and Bianchi, D. W. (2014). Circulating fetal cell-free dna fractions differ in autosomal aneuploidies and monosomy x. *Clinical chemistry*, **60**(1), 243–250.
- Schnert, A. J., Rhee, B., Comstock, D., de Feo, E., Heilek, G., Burke, J., and Rava, R. P. (2011). Optimal detection of fetal chromosomal abnormalities by massively parallel dna sequencing of cell-free fetal dna from maternal blood. *Clinical chemistry*, **57**(7), 1042–1049.
- Shubina, J., Trofimov, D. Y., Barkov, I. Y., Stupko, O. K., Goltsov, A. Y., Mukosey, I. S., Tetrushvili, N. K., Kim, L. V., Bakharev, V. A., Karetnikova, N. A., et al. (2017). In silico size selection is effective in reducing false positive nips cases of monosomy x that are due to maternal mosaic monosomy x. *Prenatal diagnosis*, **37**(13), 1305–1310.
- Snyder, M. W., Simmons, L. E., Kitzman, J. O., Coe, B. P., Henson, J. M., Daza, R. M., Eichler, E. E., Shendure, J., and Gammill, H. S. (2015). Copy-number variation and false positive prenatal aneuploidy screening results. *New England Journal of Medicine*, **372**(17), 1639–1645.
- Straver, R., Sistermans, E. A., Holstege, H., Visser, A., Oudejans, C. B., and Reinders, M. J. (2013). Wisecondor: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. *Nucleic acids research*, **42**(5), e31–e31.
- Sun, K., Chan, K. A., Hudecova, I., Chiu, R. W., Lo, Y. D., and Jiang, P. (2017). Coffee: control-free noninvasive fetal chromosomal examination using maternal plasma dna. *Prenatal diagnosis*, **37**(4), 336–340.
- Sung, K. K., Hannum, G., Geis, J., Tynan, J., Hogg, G., Zhao, C., Jensen, T. J., Mazloom, A. R., Oeth, P., Ehrich, M., et al. (2015). Determination of fetal dna fraction from the plasma of pregnant women using sequence read counts. *Prenatal diagnosis*, **35**(8), 810–815.
- Tynan, J., Kim, S., Mazloom, A., Zhao, C., McLennan, G., Tim, R., Liu, L., Hannum, G., Hull, A., Bombard, A., et al. (2016). Application of risk score analysis to low-coverage whole genome sequencing data for the noninvasive detection of trisomy 21, trisomy 18, and trisomy 13. *Prenatal diagnosis*, **36**(1), 56–62.
- Wang, L., Meng, Q., Tang, X., Yin, T., Zhang, J., Yang, S., Wang, X., Wu, H., Shi, Q., Jenkins, E. C., et al. (2015a). Maternal mosaicism of sex chromosome causes discordant sex chromosomal aneuploidies associated with noninvasive prenatal testing. *Taiwanese Journal of Obstetrics and Gynecology*, **54**(5), 527–531.
- Wang, S., Huang, S., Ma, L., Liang, L., Zhang, J., Zhang, J., and Cram, D. S. (2015b). Maternal x chromosome copy number variations are associated with discordant fetal sex chromosome aneuploidies detected by noninvasive prenatal testing. *Clinica Chimica Acta*, **444**, 113–116.
- Wang, Y., Chen, Y., Tian, F., Zhang, J., Song, Z., Wu, Y., Han, X., Hu, W., Ma, D., Cram, D., et al. (2014). Maternal mosaicism is a significant contributor to discordant sex chromosomal aneuploidies associated with noninvasive prenatal testing. *Clinical chemistry*, **60**(1), 251–259.
- Yu, C. S., Chan, K. A., Zheng, Y. W., Jiang, P., Liao, G. J., Sun, H., Akolekar, R., Leung, T. Y., Go, A. T., van Vugt, J. M., et al. (2014). Size-based molecular diagnostics using plasma dna for noninvasive prenatal testing. *Proceedings of the National Academy of Sciences*, **111**(23), 8583–8588.



Zhang, L., Zhu, Q., Wang, H., and Liu, S. (2017). Count-based size-correction analysis of maternal plasma dna for improved noninvasive prenatal detection of fetal trisomies 13, 18, and 21. *American journal of translational research*, **9**(7), 3469.

Zhao, C., Tynan, J., Ehrich, M., Hannum, G., McCullough, R., Saldivar, J.-S., Oeth, P., van den Boom, D., and Deciu, C. (2015). Detection of fetal subchromosomal abnormalities by sequencing circulating cell-free dna from maternal plasma.

*Clinical chemistry*, **61**(4), 608–616.

Zhou, X., Sui, L., Xu, Y., Song, Y., Qi, Q., Zhang, J., Zhu, H., Sun, H., Tian, F., Xu, M., et al. (2017). Contribution of maternal copy number variations to false-positive fetal trisomies detected by noninvasive prenatal testing. *Prenatal diagnosis*, **37**(4), 318–322.