medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

BinDel: software tool for detecting clinically significant microdeletions in low-coverage WGS-based NIPT samples

Priit Paluoja[1, 2], Hindrek Teder[2, 3], Andres Salumets[2, 3, 4], Kaarel Krjutškov[2, 3], Priit Palta[2, 5, 6, 7]

[1] Institute of Clinical Medicine, University of Tartu, Tartu, Estonia
[2] Competence Centre for Health Technologies, Tartu, Estonia
[3] Department of Obstetrics and Gynaecology, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia
[4] Division of Obstetrics and Gynecology, Department of Clinical Science, Intervention and Technology, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden
[5] Institute of Genomics, University of Tartu; Tartu, Estonia
[6] Activate Health, Tallinn, Estonia
[7] Institute for Molecular Medicine Finland (FIMM), University of Helsinki; Helsinki, Finland

Abstract

Clinically pathogenic chromosomal microdeletions causing genetic disorders such as DiGeorge syndrome are rare genetic aberrations that can cause clinically relevant fetal and childhood developmental deficiencies. Clinical severity of such deficiencies depend on the exact genomic location and genes affected by the fetal chromosomal aberration.

Here we present the BinDel, a novel region-aware microdeletion detection software package developed to infer clinically relevant microdeletion risk in low-coverage whole-genome sequencing NIPT data. To test BinDel, we quantified the impact of sequencing coverage, fetal DNA fraction, and region length on microdeletion risk detection accuracy. We also estimated BinDel accuracy on known microdeletion samples and clinically validated aneuploidy samples.

BinDel identified each positive control sample as high risk. We also determined that it is critical to take into account that the sample with a detected high microdeletion risk does not have a full chromosome aneuploidy, as the latter can cause erroneous high microdeletion risk findings. We observed that lower sequencing coverage resulted in reduced microdeletion detection accuracy, and higher fetal fractions considerably increased the microdeletion detection accuracy, with coverage becoming increasingly relevant as fetal DNA fraction decreased.

In conclusion, we developed an R package-based software tool BinDel for inferring fetal microdeletion risks, which accurately identified all positive control samples with microdeletion or -duplication aberrations as high-risk samples.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Keywords: microdeletions, software, NIPT, fetal fraction, sequencing coverage, microdeletion detection

Corresponding author: Priit Palta (priit.palta@gmail.com)

Introduction

Non-Invasive Prenatal genetic Testing (NIPT) is a circulating cell-free DNA (cfDNA) and sequencing-based screening test to detect the risk of fetal aneuploidies. In addition to 'traditional' screening for most common aneuploidies (causing, e.g. Down, Edwards, and Patau syndromes), whole-genome sequencing (WGS) NIPT assay data can enable the analysis of any region of the genome to infer the risk of potential subchromosomal aberrations – microdeletions. The most frequent (1:1,524) microdeletion (MD) is the 22q11.2 heterozygous deletion (2.1Mb in length), causing the DiGeorge syndrome, characterised by adverse infant and childhood outcomes [1]. Other clinically relevant microdeletions are less frequent but also cause severe genetic disorders such as, for example, Williams-Beuren (5.2Mb in length), Wolf-Hirschhorn (4.5Mb) and Prader-Willi/Angelman (PWS/AS) syndrome (5.5Mb).

While pathogenic microdeletions are individually rare, with frequencies ranging from 1:1,524 for MD causing DiGeorge syndrome to 1:50,000 for MD causing Cri-du-chat syndrome, the overall prevalence rate of microdeletions is estimated to be approximately 1:2,500 [1–3]. 1:2,500 translates to that in a population of pregnant women younger than 35 years fetal microdeletions are more common than Down syndrome [2]. As there are no biomarker-based screening tests for MDs for first trimester screening (ultrasound sonography procedure can only sometimes detect associated findings for specific MD syndromes), a reliable NIPT-based microdeletion screening is required [4]. With the advance in non-invasive prenatal testing (NIPT) algorithms and laboratory methods, the detection of such pathogenic chromosomal microdeletions during routine NIPT is starting to be incorporated into clinical settings [5–7].

Previous low-coverage WGS based NIPT studies have demonstrated that one of the most important factors of successful NIPT analysis is sequencing coverage (depth) [8]. For full chromosome aneuploidy detection, WisecondorX [8], for example, has shown to accurately operate with coverage of 5 million (M) reads per sample (RPS), while some other tools require coverage of 20M RPS to keep the number of false-negative trisomy calls under clinically acceptable 1% screening test limit [9]. The higher the sequencing coverage, the greater the likelihood of comprehensive and confident genome-wide interrogation and, therefore, more accurate evidence to infer chromosomal or even sub-chromosomal aberrations

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

[9]. Kucharik *et al*. suggested that for MD NIPT assays, sequencing coverage should be 10-20M RPS, with the exact minimum coverage depending on the fetal DNA fraction and target region length [10].

Fetal DNA fraction (FF) estimates the proportion of cell-free fetal DNA molecules originating from the fetus/placenta as compared to the total cell-free DNA molecules captured from a pregnant women's blood sample. The higher the FF, the greater the possibility of detecting potential fetal aberrations from the maternal 'background' with high confidence. For that reason, the quality control related FF cut-off threshold in complete chromosome aneuploidy detection is commonly set to 4% [11,12].

As both higher sequencing coverage and higher FF allow more confident detection of fetal aberrations, the same principle also applies to the MDs region length [10,13]. The longer the region, the more sequencing reads are expected to cover it, increasing the potential to detect the deficit or excess of sequencing reads in the studied sample and chromosomal region, consequently drawing attention to a potential MD risk.

Even though there are some existing software tools that besides full chromosomal analuploidies can detect putative MDs (e.g. WisecondorX), these tools rely on sequential fixed-length genomic bins without possibility to transfer known MD region specific context (customising bin length, removing bins or using variable bin length) into the MD analysis [8]. Furthermore, to our knowledge, these tools have not been comprehensively evaluated for detecting MDs.

Here we present BinDel, a novel microdeletion detection software package developed to infer clinically relevant microdeletion risk from low-coverage WGS-based NIPT data. To demonstrate its feasibility, we quantify the impact of sequencing coverage, FF, and region length on microdeletion detection accuracy. While screening of pathogenic microdeletions in NIPT has been studied, a systematic review by Zaninović *et al*. notes that there is no consistency in achieved sensitivity, specificity, and PPV rates depending on the copy number variation (CNV) size between different studies [2]. Here we systematically investigated the importance of chromosomal region-specific characteristics with algorithm parameters in detecting microdeletions and exact connections between FF, coverage and different MD regions. Finally, we estimate BinDel accuracy on known microdeletion and aneuploidy samples.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Results

BinDel is an R software package developed to infer microdeletion and -duplication risk in user-defined regions of interest in WGS-based NIPT samples. Given a human reference assembly aligned NIPT sample, BinDel divides sample sequencing reads into genomic bins for each user-defined microdeletion region and counts the number of reads in each bin (**Suppl. Figure 1**). Next, BinDel performs multiple normalisation steps of these data, creates informative features from the normalised read counts, calculates Mahalanobis distance from the euploid pregnancy reference group samples, and transforms the resulting distances into MD risk probabilities that are reported as output.

Here we used WGS data for 373 euploid NIPT samples (previously determined to be normal euploid pregnancies using the NIPT screening test and postnatal examination) with ~10 million (M) reads per sample (RPS) to develop and optimise the algorithm. In these samples (sample groups 1-2 in **Suppl. Table 1**), we bioinformatically reduced read count proportionally in the MD region to *in silico* induce clinically relevant fetal heterozygous MD events for various fetal DNA fractions (FF) and sequencing coverages (sample groups 9-11 in **Suppl. Table 1**), which were then blindly studied with BinDel.

Sequencing read coverage and microdeletion detection accuracy

First, we investigated the impact of sequencing coverage on microdeletion detection accuracy by applying BinDel on 73 euploid and 73 samples with *in silico* created MDs over nine pre-defined clinically relevant MD regions (**Suppl. Table 2**, sample group 2 and 10 in **Suppl. Table 1**). In addition to initial sequencing coverage (10M RPS), we also used *in silico* down- and upsampled (5M and 20M RPS, respectively) data (see Materials and Methods, sample groups 7, 8, 9 and 11 in **Suppl. Table 1**).

As expected, we observed that lower sequencing coverage resulted in reduced MD detection accuracy (**Figure 1**, **Suppl. Figure 2**, **Suppl. Figure 3**). This was observed for most studied MD regions, and in some regions, sequencing coverage had considerable effect on MD detection accuracy. For example, at 5% FF, if sequencing coverage was increased from 5M RPS to 20M, the sensitivity of microdeletion detection in Prader-Willi/Angelman (PWS/AS) associated region increased from 0.55 to 0.86 and in DiGeorge MD region, sensitivity increased from 0.18 to 0.42 (**Figure 1**F,B).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

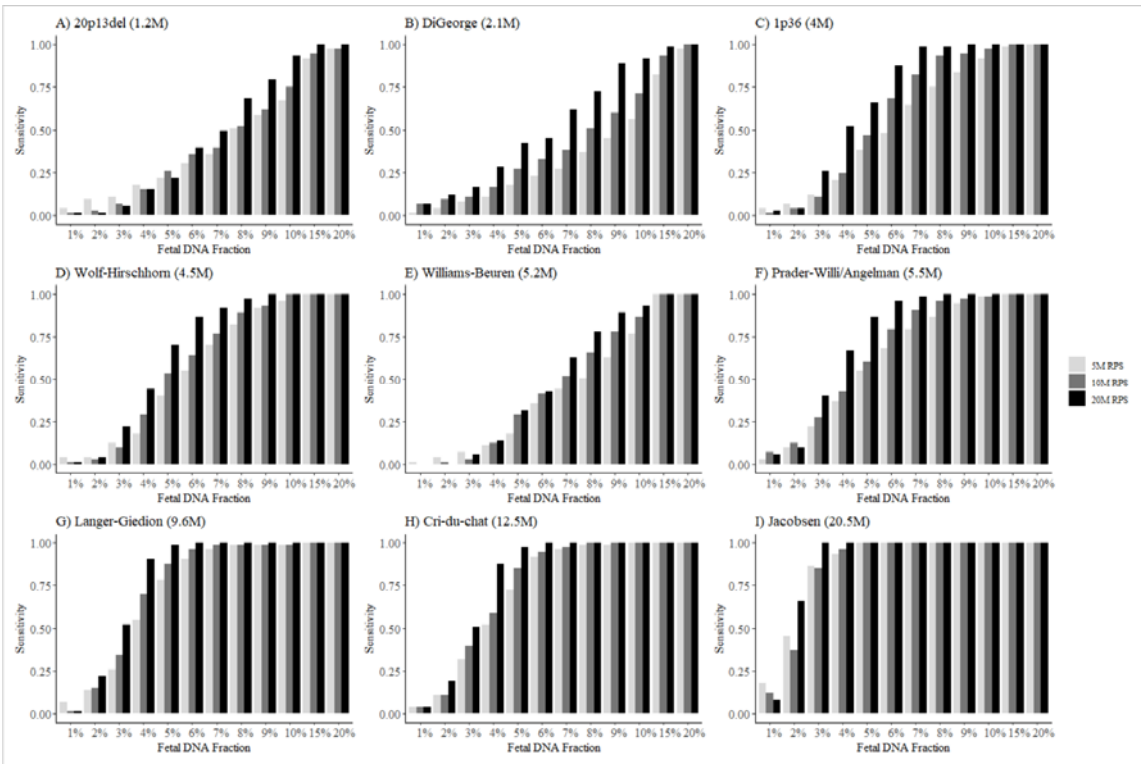


Figure 1. BinDel detection sensitivity for in silico created microdeletions. Sensitivity estimates are grey, dark grey, and black for 5, 10, and 20M reads per sample (RPS), respectively. Sensitivity calculated for 20p13del (A), DiGeorge (B), 1p36 (C), Wolf-Hirschhorn (D), Williams-Beuren (E), Prader-Willi/Angelman (F), Langer-Giedion (G), Cri-du-chat (H), and Jacobsen (I) associated MD.

Fetal DNA fraction and microdeletion detection accuracy

Next, we investigated the impact of FF on MD detection accuracy. As anticipated, we determined that fetal DNA fraction had a significant effect on MD detection accuracy (**Suppl. Figure 2**). Higher FF considerably increased the MD detection sensitivity and positive predictive value (PPV) (**Figure 1**, **Suppl. Figure 2**). For example, in the case of 20M RPS, as FF increased from 5% to 10%, microdeletion detection sensitivity in the DiGeorge syndrome associated MD region increased from 0.42 to 0.92 (**Figure 1**B). Similarly, in the case of 10M RPS and increased FF (from 5% to 10%), DiGeorge region MD detection sensitivity increased from 0.27 to 0.71 (**Figure 1**B).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Microdeletion region length and detection accuracy

We also observed that the MD chromosomal region length had a considerable impact on its detection accuracy. In the longer MD regions, the detection of *in silico* created MDs was considerably more accurate (**Figure 2**B). For example, at 8% FF ordered by region length (**Suppl. Table 2**), 20p13del (1.2Mb), DiGeorge (2.1Mb), 1p36 (4Mb), Wolf-Hirschhorn (4.5Mb), Prader-Willi/Angelman (5.5Mb), Langer-Giedion (9.6Mb), Cri-du-chat (12.5Mb) and Jacobsen region (20.5Mb) MD detection sensitivity increased together with the MD region length (sample group 10, **Figure 2**B). The only exception to this was observed in Williams-Beuren (5.2Mb) MD region, which had lower sensitivity than other similar-length MD regions (**Figure 2**B).

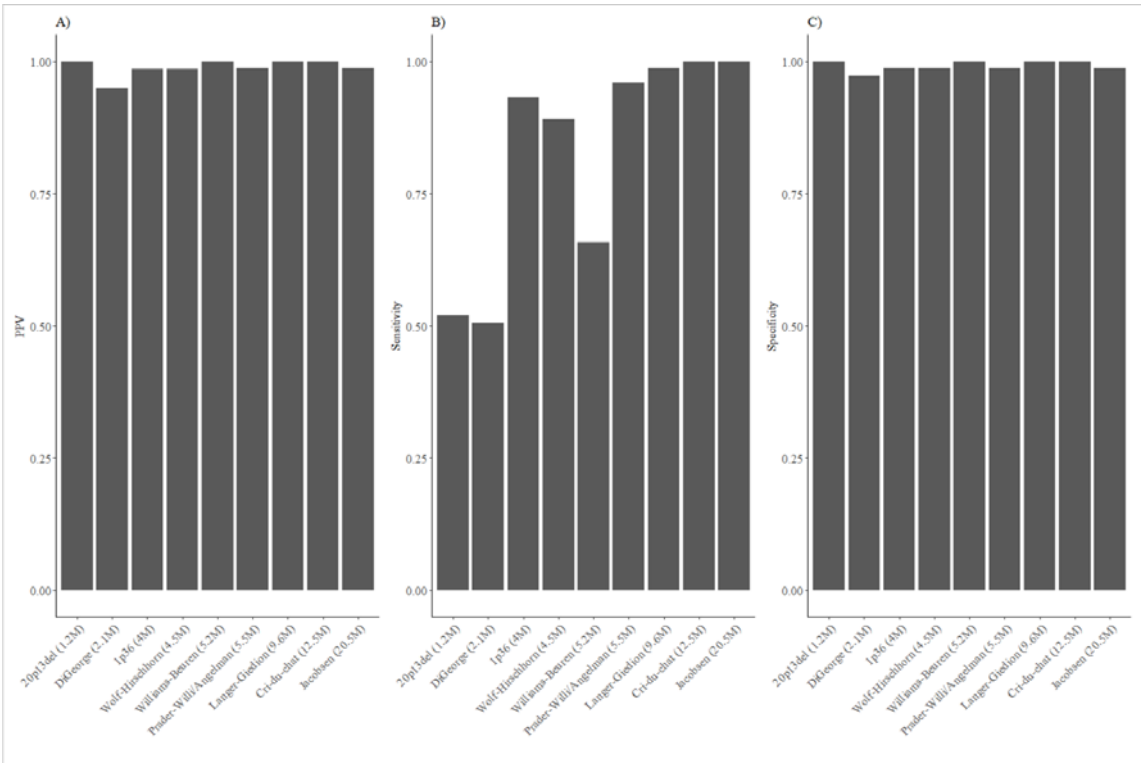


Figure 2. Microdeletion associated region length effect on the accuracy of detecting microdeletions. Positive predictive value (A), sensitivity (B) and specificity (C) were calculated for in silico created microdeletions at 10M reads per sample (RPS) with FF of 8%.

Region-specific characteristics in MD detection

Analysis of the MD regions in WGS NIPT should consider local chromosomal characteristics as MD regions are often in the vicinity of low-copy repeat (LCR) or homologous sequences, which in the case of short read sequencing methods do not facilitate unambiguously unique sequencing read mapping [14]. Our context-aware region analysis demonstrates that if each

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

MD region specific characteristics are considered (such as region length, region-specific bin size and regional variability) the accuracy of detecting corresponding potential MDs can improve (**Suppl. Figure 7**). A good example is the PWS/AS syndrome associated region, where we examined the normalised read count distribution of the BinDel 10M RPS reference set (300 samples, sample group 1 in **Suppl. Table 1**) and determined two highly variable 300 kb bins (**Suppl. Figure 5**). After removing these bins from the analysis, we observed a slightly increased sensitivity at FF of 3% and 5% (**Suppl. Figure 6**B, F). Determining and removing variable bins could be performed for any MD region under analysis. For example, BinDel's reference set file contains read counts for each bin (per sample) in the plain text format. It is easy to visualise the reference data for each MD region, determine the ultra-variable bins and remove them from the reference file.

BinDel accuracy on validated microdeletion samples

To evaluate the accuracy of our algorithm with the optimal selected settings, we analysed six low-coverage (5 to 20 M RPS) WGS-based NIPT assays with confirmed fetal aberration events. These included two confirmed DiGeorge syndrome MD samples and one sample possessing confirmed 20p13 microdeletion and 3q29 microduplication events (sample groups 4-6 in **Suppl. Table 1**).

BinDel successfully inferred the high risk for correct regions (DiGeorge, 3q29, and 20p13) in all validation samples (**Figure 3**A-F). BinDel also erroneously inferred one false positive high-risk MD in the DiGeorge MD region for sample SC-0106, sequenced at 5M RPS (**Figure 3**F, DiGeorge region), emphasising the relevance of sufficiently high sequencing coverage (≥10M RPS), which would enable avoiding false positive MD findings in clinical applications. BinDel's MD detection sensitivity, specificity and PPV on validated MD samples are shown in **Table 1** (10 and 20M RPS).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

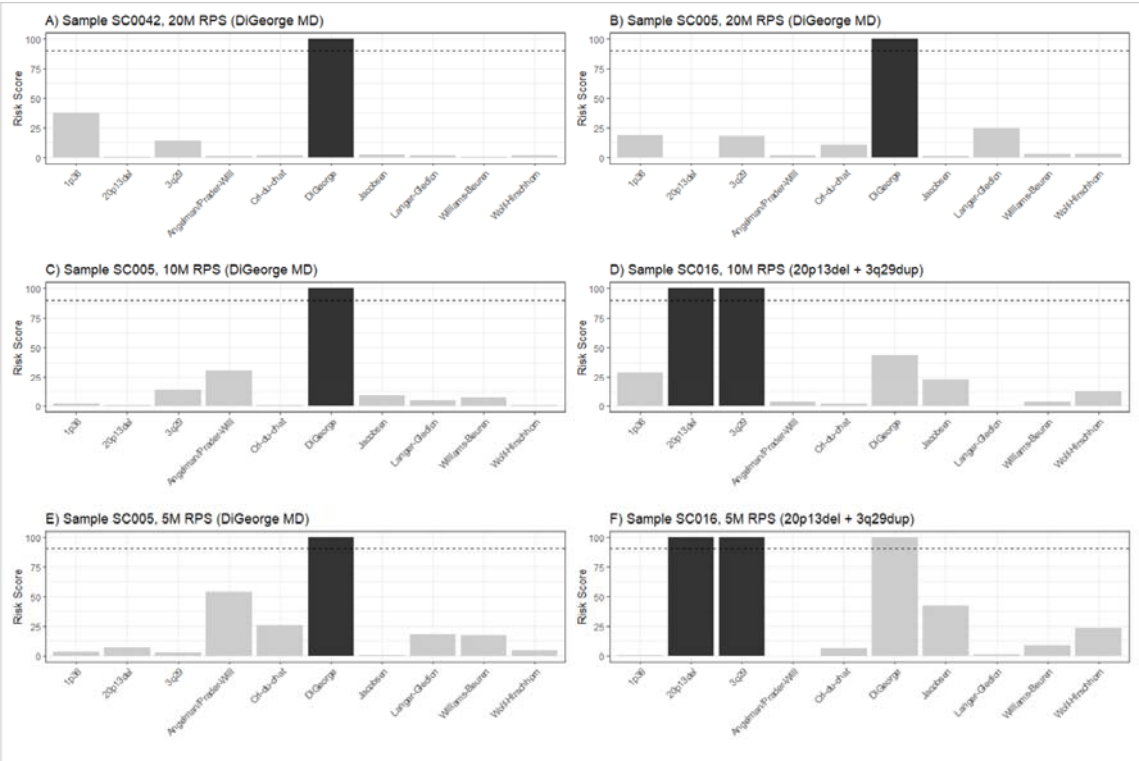


Figure 3. BinDel inferred risk scores for samples with microdeletion or -duplication syndromes across a range of sequencing depths. The regions affected by the syndrome are shown in black. The dashed line represents the cut-off point for high-risk calls in each graph. Estimates were obtained at sequencing depths of 20M reads per sample (RPS) (A, B), 10M RPS (C, D), and 5M RPS (E, F).

Table 1. BinDel's microdeletion detection sensitivity, specificity and positive predictive value (PPV) for 20p13, Prader-Willi and DiGeorge microdeletion region. Metrics were calculated over sample group 3 (euploid and Prader-Willi only) combined with sample groups 4 and 5.

Microdeletion region	Sensitivity	Specificity	PPV
20p13 deletion	1	0.991	0.333
Prader-Willi	1	0.974	0.143
DiGeorge	1	0.987	0.25

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

BinDel accuracy on clinically validated aneuploidies

We further analysed 244 clinically validated samples (sequenced at 20M RPS, sample group 3 in **Suppl. Table 1**), including 17 full chromosome fetal aneuploidies and one fetal Prader-Willi syndrome associated MD. BinDel successfully detected the confirmed Prader-Willi syndrome case (**Figure 4**C). BinDel incorrectly detected high PWS/AS risk for seven samples lacking PWS/AS aberration, resulting in a PPV of 12.5% for PWS/AS (95% CI 0–35%), one of which was a trisomy 21 (T21) sample. Besides one T21 sample with false-positive PWS/AS MD risk, we observed three more aneuploidy samples with false-positive MD risk - two in Wolf-Hirschhorn MD region (samples with T21 ad T18) and one in Jacobsen syndrome associated MD region (sample with T21). Consequently, if MD risk is detected, a sample should also be screened for aneuploidies, as existing full chromosome aneuploidy can cause false MD risk indications. Although BinDel is not specifically developed for chromosomal aneuploidy detection, it was able to detect 1/1 T13 (PPV 17%, 95% CI 0–46%), 5/5 T18 (PPV 83%, 95% CI 54–100%), and 11/12 T21 (PPV of 85%, 95% CI 65–100%) in known trisomy samples (**Figure 4**A). BinDel falsely detected 20 (8.2%) high risk aneuploidies and 36 high risk MDs (14.8%, four of which were known trisomy samples) (**Figure 4**A, **Figure 4**C). This is fairly similar to the false positive proportions (9.6% for MDs and 15% for aneuploidies) observed in the 10M RPS sample group 2 (**Suppl. Table 1**) that used in the algorithm development and optimisation process (**Figure 4**B, **Figure 4**D).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

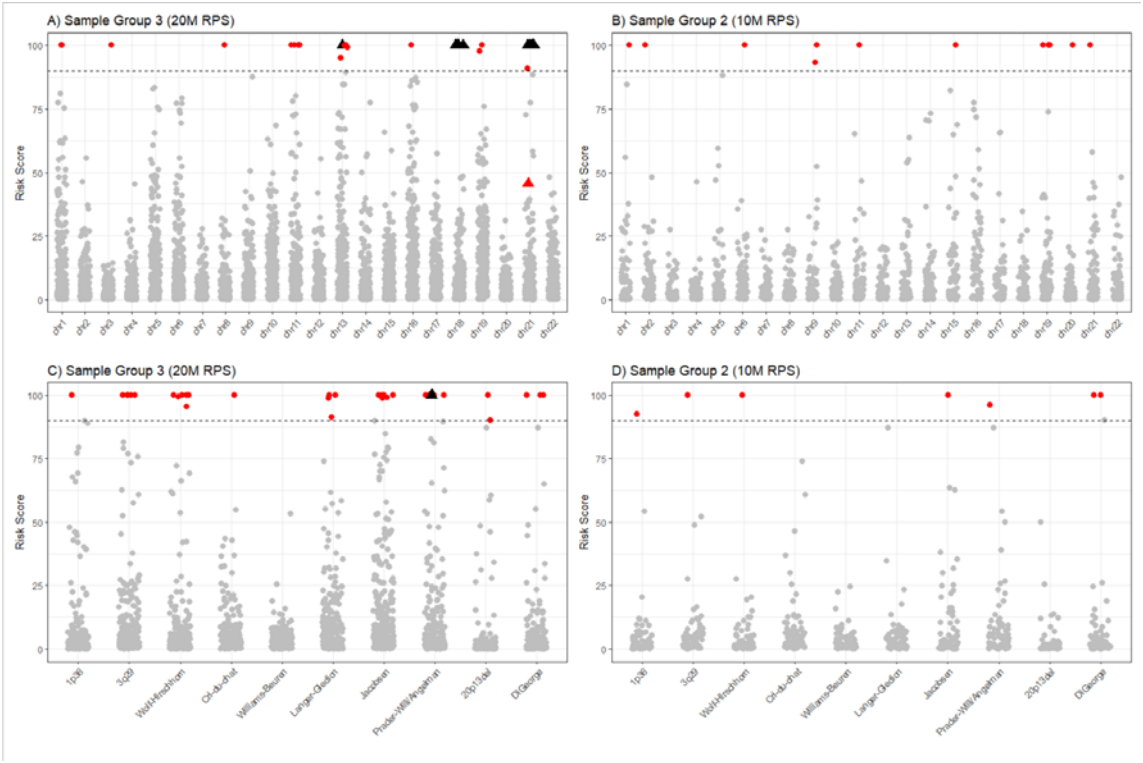


Figure 4. BinDel inferred risk scores for samples sequenced on 10 and 20M reads per sample (RPS). Black triangles show detected cases, undiscovered cases are represented by red triangles, and red circles represent euploid samples with wrongly inferred high-risk. The dashed line represents the cut-off point for high-risk calls in each graph. Data obtained with sequencing depths of 20M RPS (A), 10M RPS (B), 20M RPS (C), and 10M RPS (D) are shown.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Discussion

NIPT is a fetal aneuploidy risk screening method that is based on laboratory and computational analysis of cell-free DNA extracted from pregnant women's peripheral blood. NIPT can also be used to screen for rare, but clinically relevant pathogenic chromosomal microdeletions. The most frequent (1:1,524) microdeletion is the 22q11.2 heterozygous deletion (2.1Mb in length, **Suppl. Table 2**), causing adverse infant and childhood outcomes (DiGeorge syndrome) [1]. Other clinically significant microdeletions are less frequent but also cause severe genetic disorders (**Suppl. Table 2**).

Here we present BinDel, a novel microdeletion detection software package to infer clinically significant microdeletion risk from low-coverage WGS NIPT data. To demonstrate its feasibility, we quantify the impact of sequencing coverage, FF, and region length on microdeletion detection accuracy. We also estimated BinDel accuracy on known microdeletion samples and clinically validated aneuploidy samples.

In MD regions longer or equal to 4.5Mb in length, we discovered that FF is one of the most important factors for correct MD detection. MDs of 4.5Mb and longer were determined with 100% sensitivity from FF of 15% onwards at all coverages measured (**Figure 1D-I**). Coverage and FF were equally important components for MD detection sensitivity in regions shorter than 4.5Mb, with the coverage impact on sensitivity significantly reducing with an FF of 20% (**Figure 1A-C**). Specificity was only affected by the coverage in all MD regions analysed (**Suppl. Figure 4**). Our analysis demonstrated that considering region-specific characteristics such as region length, region-specific bin size, and regional variability can improve the accuracy of detecting corresponding potential MDs (**Suppl. Figure 6**, **Suppl. Figure 7**). In microdeletion detection accuracy analyses, BinDel – an R package based software tool for inferring fetal MD risks – accurately identified all positive control samples (sample groups 4-6 in **Suppl. Table 1**) with microdeletion or -duplication aberrations as high risk samples (**Figure 3**A-F). Finally, each analysed sample should also be screened for aneuploidies, as existing aneuploidy can cause false high-risk MD predictions, as we observed in four clinically validated trisomy samples. We believe that the number of false positive high-risk MD results can be reduced in a clinical context by confirming that there is no fetal full chromosome aneuploidy detected in the same sample, followed by resequencing and reanalysis of the sample to avoid the random sequencing read placement affecting the genomic site of the MD. Although not explicitly developed for full chromosome aneuploidy

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

detection, BinDel also accurately identified 17 of the 18 aneuploidies as high-risk results (**Figure 4**A).

Dar *et al.* have shown recently that NIPT screening for DiGeorge syndrome can detect most cases, with a PPV of 23.7% (95% CI, 11.44–40.24%) [1]. Wang *et al.* observed PPV of 49.0% for fetal microdeletion/microduplication syndromes [15]. Due to the different sensitivity of the algorithms and the influence of random sequencing read alignment effects, such PPVs, as seen by Dar *et al.* and Wang *et al.*, are in accord with our hypothesis that PPV might increase if inferred high-risk samples were resequenced [1,15]. Rafalko *et al.* associated an overall PPV of 74.2% for detecting sub-chromosomal copy number variations [16], which is higher than BinDel's PPV on a very limited number of validated microdeletion samples (**Table 1**) but generally lower than BinDel's PPV on *in silico* created MD samples (**Suppl. Figure 2**).

While our study included a significant number of NIPT samples to create a representative reference panel, we only had a small number of confirmed microdeletion samples. We also observed a considerable number of false positive results for MDs (14.8%, **Figure 4**C) in the analysis of 20M RPS of 244 clinically validated samples (**Suppl. Table 1**, sample group 3). In practice, we believe that such high-risk samples should be resequenced at higher coverage to reduce the number of false positive MD results. Further, we analysed different datasets with a single 10M RPS development reference set. Results would be more consistent if the reference matched the analysis data set coverage more accurately, as demonstrated previously for full chromosome aneuploidy detection [9]. Using a single 10M RPS development reference set (sample group 1 in **Suppl. Table 1**) to analyse various laboratory protocols, on the other hand, demostrates that BinDel can handle different NIPT protocols. Second, in the development phase, upsampling 10M to 20M RPS (sample group 8 in **Suppl. Table 1**) may not perfectly reflect actual sample sequencing read variability since it might reduce or increase observed sequencing read data variability. Moreover, compared to the *in silico* MDs, naturally occurring MDs would provide a more accurate measure but would necessitate a relatively large sample set of rare microdeletions based on a single laboratory protocol, which is difficult to obtain.

Finally, for PWS/AS region analysis, we noticed that MD detection accuracy can be increased at low FF by leaving out the ultra-variable bins (**Suppl. Figure 6**B, C), but we also saw no effect at FF of 1%, 10% and 20%, and a small decrease at 8% (**Suppl. Figure 6**A, E, F, D

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

respectively), which could be linked to the relatively long bin length (300bp) used in these analyses.

In conclusion, we developed a software tool for MD detection in WGS-NIPT and evaluated it with a number of positive and negative control samples, whilst also considering microdeletion detection accuracy in terms of fetal DNA fraction, sequencing coverage and MD region length.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Materials and Methods

Ethics statement

This study was performed with the written informed consent from the participants and with approval of the Research Ethics Committee of the University of Tartu (#352/M-12).

Studied samples and sample pre-processing

Multiple NIPT datasets were used for this study, including development, validation, DNA mixes and samples with in silico created MDs (**Suppl. Table 1**). We had separate dataset(s) for the BinDel reference set, BinDel development, MD validation, algorithm specificity, PPV and sensitivity analysis. A selection of datasets was sub- or upsampled, or MD was bioinformatically created in silico (as described below).

Each sequenced sample was aligned against human reference assembly GRCh38, sorted, and the reads originating from a single fragment of DNA were marked as duplicates.

Sample group 1 (**Suppl. Table 1**) consisted of 300 samples reported previously as euploid fetus pregnancies by NIPTIFY screening test and postnatal evaluation. The group consisted of 151 female and 149 male fetus pregnancies (determined by NIPT) with coverage between 8.0-18.5M RPS (median of 12.7M) and was used as a BinDel 10M RPS euploid reference set. Sample group 1 was processed similarly to previously published guidelines from KU Leuven, with modifications [17]. Briefly, peripheral blood samples were collected in cell-free DNA BCT tubes (Streck, USA), and plasma was separated with standard dual centrifugation. Cell-free DNA was extracted from 3 ml plasma using MagMAX Cell-Free DNA Isolation Kit (ThermoFisher Scientific). Whole-genome libraries were prepared using the FOCUS (Fragmented DNA Compact Sequencing Assay, Competence Centre on Health Technologies, Estonia) NIPT method protocol with 12 cycles for the final PCR enrichment step. In the following quantification, equal amounts of 36 samples were pooled, and the quality and quantity of the pool were assessed on Agilent 2200 TapeStation (Agilent Technologies, USA). Whole genome sequencing was performed on the NextSeq 550 instrument (Illumina Inc.) with an average coverage of 0.32× (minimum 0.08 and maximum 0.42) and producing 85 bp single-end reads.

Sample group 2 consisted of 73 samples used for BinDel development with coverage between 1.9-17.4M (median 12.9) RPS and was processed as was sample group 1. Similarly to sample group 1, sample group 2 samples were reported previously as euploid fetus pregnancies by

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

NIPTIFY screening test and postnatal evaluation. The sample set consisted of 39 female and 28 male fetus pregnancies. For six samples, the gender of the fetus was not determined previously by NIPT due to the patient's wish. This set also consisted of one SeraCare Life Sciences Inc euploid fetus reference material as a negative control sample. Sample group 2 was used as a development dataset of negative controls when calculating PPV, sensitivity, and specificity in combination with sample group 10. Sample group 2 was also used for in silico introduction of MDs in sample group 10 and for up- and downsampling samples in sample groups 7 and 8.

Sample group 3 is based on a previously published validation study by Žilina *et al.* [12] (except for one Prader-Willi sample, which was not in the previous study) and consists of 244 clinically validated samples. These samples include one T13, five T18, 12 T21, one Prader-Willi and 225 euploid fetus samples. Sample group 3 was sequenced with Illumina NextSeq 500 platform, producing 85 bp single-end reads with an average coverage of 0.32× at the University of Tartu, Institute of Genomics Core Facility, as described previously by Žilina *et al.* [12]. The sample group was used for algorithm validation as a test sample set.

SC005 (SeraCare Life Sciences Inc lot #10446565), SC0042 (#10571706), and SC016 (#10560229) are SeraCare Life Sciences Inc circulating cell-free DNA (ccfDNA) like mixture of human genomic DNA that consists of matched maternal and fetus DNA (sample groups 4-6 in **Suppl. Table 1**). SC005 (male fetus pregnancy) and SC0042 (female fetus pregnancy) consist of matched DNA of maternal and fetus with DiGeorge Syndrome. SC005 DiGeorge MD is at least 2Mb and includes TUPLE1 deletion. SC016 is a custom-ordered DNA Mix from SeraCare Life Sciences Inc with fetus DNA having a pathogenic loss of the terminal region of 20p13 (contains 14 genes including SOX12 and NRSN2 but not including JAG1) and a pathogenic 3q29 duplication. The specification states that the gain of 3q is at least 4.3 M in size, and the loss of 20p is at least 1.2 M. Sample group 4 is sequenced at 20M RPS, sample group 5 at 10M RPS and sample group 6 is subsampled (as is subsampled sample group 7, see below) from sample group 5. Sample groups 4-6 were used for validating MD detection.

Sample groups 7 and 8 consists of down- and upsampled samples from sample group 2 (**Suppl. Table 1**). Seventy-three euploid samples from sample group 2 were emulated to reduce and increase the RPS to 5M (by reducing read count by 50%) and 20M sequencing coverage, respectively. For 20M RPS dataset upsampling (see sample group 8 in **Suppl. Table 1**), the read count of the 73 initial samples was increased by combining initial samples

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

randomly. More precisely, combinations of 73 samples were taken twice without repetition and merged. While the intuitive would be to increase sequencing reads by repeating existing reads within the sample by the desired fraction, such increase would lead to the introduction of duplicated sequencing reads and require estimation of the sequencing read distribution in different genomic regions. From 2,628 ways to merge two samples, 73 combinations were randomly chosen to make 20M RPS analysis comparable with 10 and 5M RPS analysis. We considered that merging two different samples leads to possible copy-number variation merging, due to which we did not merge more than two samples. Merging more than two samples will eventually even out read count distribution on copy number variation loci. Sample groups 7 and 8 were used as a development dataset of negative controls (in combination with sample groups 9 and 11) when calculating PPV, sensitivity, and specificity, matching the coverage. Sample group 7 and 8 was also used for in silico introduction of MDs in sample groups 9 and 11.

Sample group 9, 10 and 11 was created in silico bioinformatically introducing heterozygous MDs by reducing sequencing reads in selected genomic regions (see **Region selection and Suppl. Table 1**) in sample groups 7, 2, 8, respectively. We introduced heterozygous microdeletions by reducing sequencing reads in selected genomic regions (see **Region selection**). Given that the NIPT sample is euploid, the FF of the sample has no effect on in silico creating and detecting microdeletions as low-coverage whole-genome sequencing NIPT does not distinguish reads between pregnant women and the fetus. We reduced the proportion of reads in the selected region in the aligned sample. For example, for creating heterozygous deletion with 10% FF, we would reduce aligned sequencing reads in the region $0.1\ (FF) \times 0.5\ (heterozygous) \times total\ sample\ read\ count$. Sample groups 9, 10, 11 combined with 7, 2, 8 (matching the coverage) was used to conduct BinDel's specificity, PPV and sensitivity analysis.

Region selection

While common pathogenic copy number variations such as 1p36 deletion are well described and defined in OMIM or DECIPHER, determining the exact start and end coordinates for the precision algorithm is not straightforward. For instance, with 1p36 deletion syndrome, OMIM and DECIPHER have non-overlapping coordinates of 23,600,000-27,600,000 and 10,001-12,780,116 (**Suppl. Table 2**) [18,19]. Furthermore, for some syndromes such as DiGeorge syndrome, OMIM did not define exact coordinates, but the DECIPHER did (**Suppl. Table 2**) [18,19]. Finally, for 20p13 microdeletion syndrome, we did not find coordinates from

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

DECIPHER or OMIM and had to use coordinates found from the SeraCare Life Sciences Inc sample specification. Given previously defined constraints, we selected OMIM coordinates over DECIPHER if both had coordinates defined, DECIPHER coordinates if OMIM did not provide coordinates, and SeraCare Life Sciences Inc custom DNA-mix sample specification coordinates if no coordinates were found from OMIM or DECIPHER.

Optimal bin size

To choose the optimal bin size and BinDel sensitivity setting, we focused on finding a bin size that would perform well over all coverages of 5, 10, and 20M RPS while also considering both false and true positive MD risk inference and would equally work well with all 9 MD regions. For this, we calculated the area under the curve (AUC) for each of the 9 MD regions with each bin size (100-500 kb), coverage (5, 10, 20M RPS), and algorithm sensitivity (greedy, conservative) setting based on the inferred BinDel risk scores (each bin size × 9 MD regions × two algorithm settings × three coverage settings, sample groups 2, 7, 8, 9, 10 and 11 in **Suppl. Table 1**). We used a fixed number of PCA components set to 270 (90% of the reference samples derived from the BinDel development phase). Next, we calculated the mean of the 9 MD AUC values per coverage and bin size (**Suppl. Figure 7**). Finally, we summed the mean AUC values over coverages for optimal bin size. We found that the 300 kb bin size with a conservative setting yields the highest sum of mean AUC over all coverages (2.693), e.g. BinDel distinguished the positive and negative MD samples best with 300 kb conservative sensitivity. From the most optimal to the least optimal bin size, the sizes were 300, 200, 100, 500, and 400 kb (sum of mean 2.679), all conservative sensitivity, followed by greedy sensitivity of 200, 100, 300, 500, and 400 (sum of mean 2.630).

BinDel algorithm

BinDel allows flexible NIPT sample analysis, where the laboratory can incorporate their knowledge about the target regions into the analysis by, for example, leaving out variable parts of the target regions or using custom varying bin lengths. BinDel outputs two risk scores, a more sensitive greedy score and a less sensitive conservative score, from which users can choose the one based on the algorithm's intended use or the region characteristics under the question. The greedy score is more sensitive but can lead to increased false-positive MD risk calls due to the increased risk of registering noise as MD risk. The conservative score is more balanced than the greedy score in distinguishing MD risks from noise. See **Optimal bin size** for more information about the bin size and score type dynamics on MD detection accuracy.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Algorithm working principles are shown in **Suppl. Figure 1**. Calculating a high-risk probability for a sample of interest requires the euploid reference sample group. Creating a file containing euploid sequencing read info by the package requires GRCh38 aligned euploid NIPT samples and pre-defined genomic bin coordinates. The reference file also defines the regions of interest.

Once the reference file is created, for GRCh38 aligned sample of interest, the genomic bins are GC% corrected [20] and normalised by total sequencing read count and bin lengths. Next, the normalised values are transformed by dividing them by dimensionality-reduced versions [21]. After the normalisation, based on the euploid reference group, an intermediate Z-score for each bin is calculated [21]. The Z-scores of each bin are aggregated and transformed so that only two values (features) describing the sample are left. The first value is received by dividing each bin Z-score by the square root of the number of bins and then summed. For the second value, each Z-score is divided by the square root of the number of bins and further divided by the count of bins over the mean reference group (region) read count and then summed. In the construction of the second value, the Laplace smoothing is applied to avoid division by zero.

After the feature engineering, the Mahalanobis distance, based on the newly created features, from the reference group is found and transformed into high-risk probabilities with Chi-Square distribution and further transformed with -log10 to human-interpretable probabilities. Conservative probabilities are calculated only based on the second engineered feature. Greedy probabilities use both Z-scores.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Acknowledgements

The authors would like to thank Signe Mölder, who participated in SeraCare Life Sciences Inc sample management and communications. This study was supported by Enterprise Estonia (grants No. EU48695 and EU53935); Estonian Research Council (grant No. PRG1076) and Horizon 2020 innovation grant (ERIN, grant no. EU952516).

References

1. Dar P, Jacobsson B, Clifton R, Egbert M, Malone F, Wapner RJ, et al. Cell-free DNA screening for prenatal detection of 22q11.2 deletion syndrome. Am J Obstet Gynecol. 2022. doi:10.1016/J.AJOG.2022.01.002

2. Zaninović L, Bašković M, Ježek D, Bojanac AK. Validity and Utility of Non-Invasive Prenatal Testing for Copy Number Variations and Microdeletions: A Systematic Review. J Clin Med. 2022;11: 3350. doi:10.3390/JCM11123350

3. Cerruti Mainardi P. Cri du Chat syndrome. Orphanet J Rare Dis. 2006;1. doi:10.1186/1750-1172-1-33

4. Lin TY, Hsieh TTA, Cheng PJ, Hung TH, Chan KS, Tsai C, et al. Taiwanese Clinical Experience with Noninvasive Prenatal Testing for DiGeorge Syndrome. Fetal Diagn Ther. 2021;48: 672–677. doi:10.1159/000519057

5. Gross SJ, Stosic M, McDonald□McGinn DM, Bassett AS, Norvez A, Dhamankar R, et al. Clinical experience with single□nucleotide polymorphism□based non□invasive prenatal screening for 22q11.2 deletion syndrome. Ultrasound Obstet Gynecol. 2016;47: 177–183. doi:10.1002/uog.15754

6. Martin K, Iyengar S, Kalyan A, Lan C, Simon AL, Stosic M, et al. Clinical experience with a single-nucleotide polymorphism-based non-invasive prenatal test for five clinically significant microdeletions. Clin Genet. 2018;93: 293–300. doi:10.1111/cge.13098

7. Helgeson J, Wardrop J, Boomer T, Almasri E, Paxton WB, Saldivar JS, et al. Clinical outcome of subchromosomal events detected by whole□genome noninvasive prenatal testing. Prenat Diagn. 2015;35: 999–1004. doi:10.1002/pd.4640

8. Raman L, Dheedene A, De Smet M, Van Dorpe J, Menten B. WisecondorX: improved copy number detection for routine shallow whole-genome sequencing. Nucleic Acids

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Res. 2018;47: 1605–1614. doi:10.1093/nar/gky1263

9. Paluoja P, Teder H, Ardeshirdavani A, Bayindir B, Vermeesch J, Salumets A, et al. Systematic evaluation of NIPT aneuploidy detection software tools with clinically validated NIPT samples. Li J, editor. PLOS Comput Biol. 2021;17: e1009684. doi:10.1371/journal.pcbi.1009684

10. Kucharik M, Gnip A, Hyblova M, Budis J, Strieskova L, Harsanyova M, et al. Non-invasive prenatal testing (NIPT) by low coverage genomic sequencing: Detection limits of screened chromosomal microdeletions. Chan KYK, editor. PLoS One. 2020;15: e0238245. doi:10.1371/journal.pone.0238245

11. Gregg AR, Skotko BG, Benkendorf JL, Monaghan KG, Bajaj K, Best RG, et al. Noninvasive prenatal screening for fetal aneuploidy, 2016 update: a position statement of the American College of Medical Genetics and Genomics. Genet Med. 2016;18: 1056–1065. doi:10.1038/GIM.2016.97

12. Žilina O, Rekker K, Kaplinski L, Sauk M, Paluoja P, Teder H, et al. Creating basis for introducing non□invasive prenatal testing in the Estonian public health setting. Prenat Diagn. 2019;39: 1262–1268. doi:10.1002/pd.5578

13. Zhao C, Tynan J, Ehrich M, Hannum G, McCullough R, Saldivar JS, et al. Detection of Fetal Subchromosomal Abnormalities by Sequencing Circulating Cell-Free DNA from Maternal Plasma. Clin Chem. 2015;61: 608–616. doi:10.1373/CLINCHEM.2014.233312

14. Watson CT, Marques-Bonet T, Sharp AJ, Mefford HC. The Genetics of Microdeletion and Microduplication Syndromes: An Update. Annu Rev Genomics Hum Genet. 2014;15: 215. doi:10.1146/ANNUREV-GENOM-091212-153408

15. Wang C, Tang J, Tong K, Huang D, Tu H, Li Q, et al. Expanding the application of non-invasive prenatal testing in the detection of foetal chromosomal copy number variations. BMC Med Genomics. 2021;14: 1–11. doi:10.1186/s12920-021-01131-6

16. Rafalko J, Soster E, Caldwell S, Almasri E, Westover T, Weinblatt V, et al. Genome-wide cell-free DNA screening: a focus on copy-number variants. Genet Med 2021 2310. 2021;23: 1847–1853. doi:10.1038/s41436-021-01227-5

17. Bayindir B, Dehaspe L, Brison N, Brady P, Ardui S, Kammoun M, et al. Noninvasive

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

prenatal testing using a novel analysis pipeline to screen for all autosomal fetal aneuploidies improves pregnancy management. Eur J Hum Genet. 2015;23: 1286–1293. doi:10.1038/ejhg.2014.282

18. Firth H V., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009;84: 524–533. doi:10.1016/J.AJHG.2009.03.010

19. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an Online catalog of human genes and genetic disorders. Nucleic Acids Res. 2015;43: D789–D798. doi:10.1093/nar/gku1205

20. Johansson LF, De Boer EN, De Weerd HA, Van Dijk F, Elferink MG, Schuring-Blom GH, et al. Novel algorithms for improved sensitivity in non-invasive prenatal testing. Sci Rep. 2017;7: 1–11. doi:10.1038/s41598-017-02031-5

21. Huijsdens–van Amsterdam K, Straver R, van Maarle MC, Knegt AC, Van Opstal D, Sleutels F, et al. Mosaic maternal 10qter deletions are associated with FRA10B expansions and may cause false-positive noninvasive prenatal screening results. Genet Med. 2018;20: 1472–1476. doi:10.1038/gim.2018.32

22. Shaffer LG, Heilstedt HA. Terminal deletion of 1p36. Lancet. 2001;358: S9. doi:10.1016/S0140-6736(01)07022-2

23. Maas NMC, Van Buggenhout G, Hannes F, Thienpont B, Sanlaville D, Kok K, et al. Genotype–phenotype correlation in 21 patients with Wolf–Hirschhorn syndrome using high resolution array comparative genome hybridisation (CGH). J Med Genet. 2008;45: 71–80. doi:10.1136/JMG.2007.052910

24. Strømme P, Bjømstad PG, Ramstad K. Prevalence estimation of Williams syndrome. J Child Neurol. 2002;17: 269–271. doi:10.1177/088307380201700406

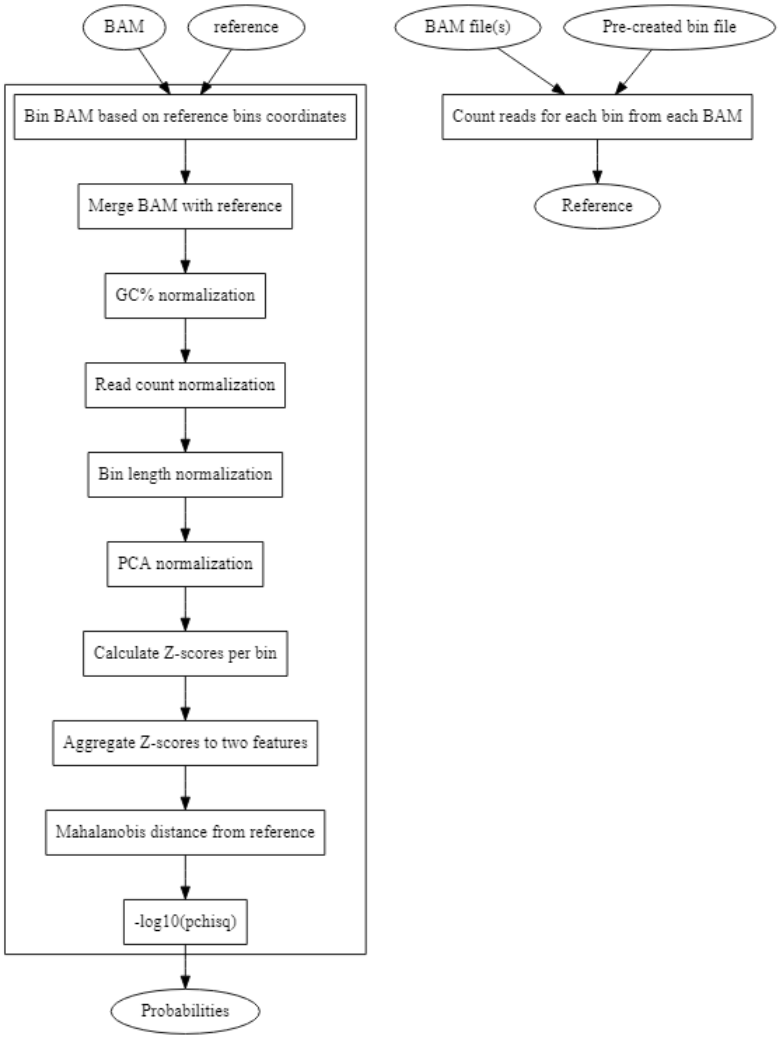
25. Ji T, Wu Y, Wang H, Wang J, Jiang Y. Diagnosis and fine mapping of a deletion in distal 11q in two Chinese patients with developmental delay. J Hum Genet 2010 558. 2010;55: 486–489. doi:10.1038/jhg.2010.51

26. Buiting K. Prader–Willi syndrome and Angelman syndrome. Am J Med Genet Part C Semin Med Genet. 2010;154C: 365–376. doi:10.1002/AJMG.C.30273

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

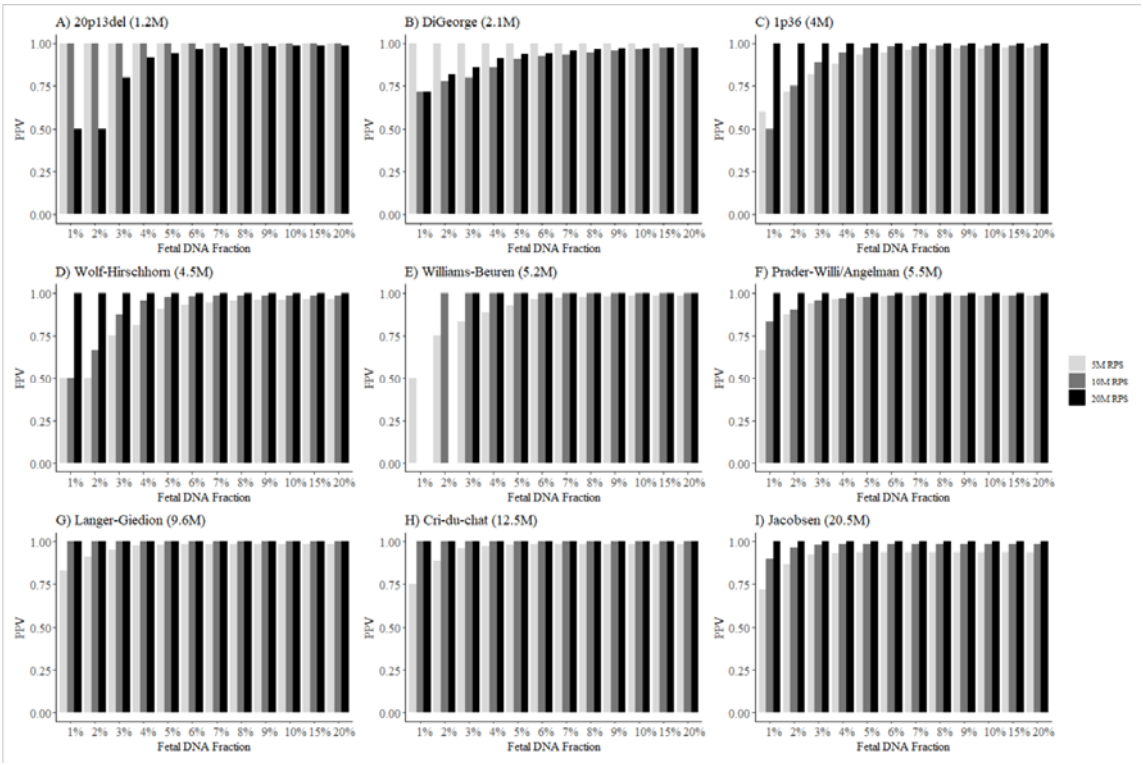
Supplementary Information



Suppl. Figure 1. Algorithm internal flow. There exist two flows. First, a reference creation process is shown on the right side of the figure. Technically, all the samples to be used as a reference set in the format of Binary Alignment Map (BAM) files are binned based on the pre-created bin coordinates description file. The output consists of binned read counts of all the used samples. The second flow, on the left of the figure, shows the process of determining MD risk on the input BAM file with the euploid fetus reference file. The input BAM reads are binned based on the reference file, and then the read counts are normalised. An intermediate Z-score for each normalised bin is calculated based on the euploid reference group. Each bin's Z-scores are aggregated so that only two features describing the sample are left. Based on the newly created features, the Mahalanobis distance from the reference group is found and transformed into MD high-risk probabilities with Chi-Square distribution and further transformed with -log10 to human-interpretable probabilities.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

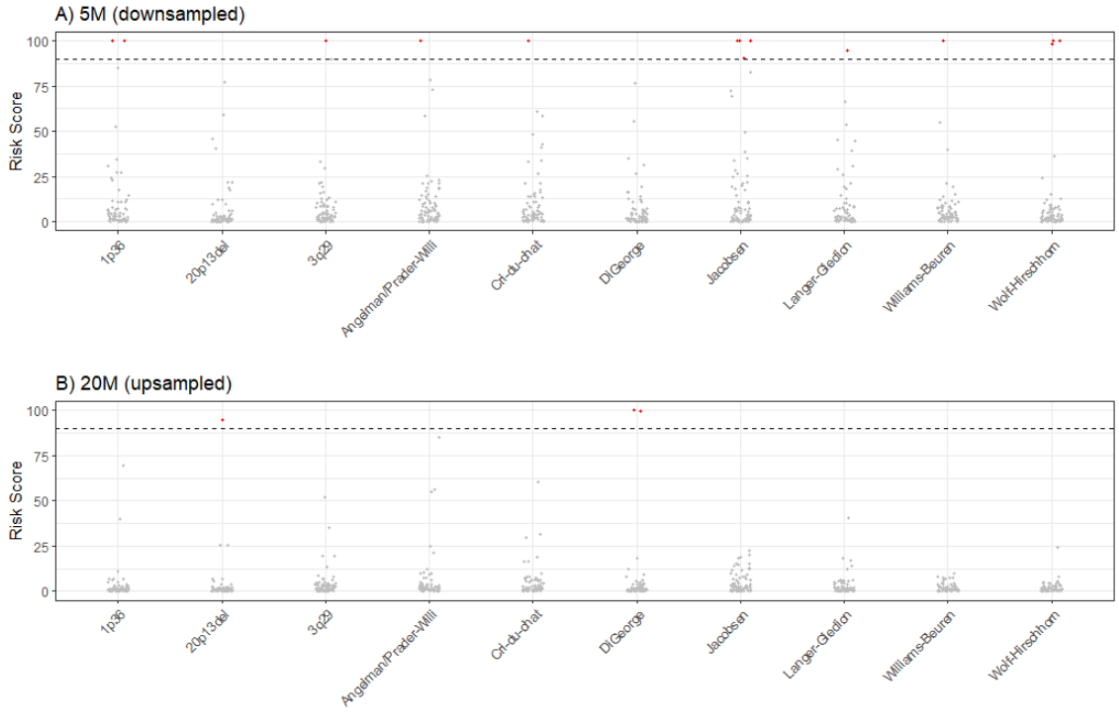
It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 2. BinDel positive predictive value (PPV) for in silico created microdeletion detection. PPV estimates are grey, dark grey, and black for 5, 10, and 20M reads per sample (RPS), respectively. PPV calculated for 20p13del (A), DiGeorge (B), 1p36 (C), Wolf-Hirschhorn (D), Williams-Beuren (E), Prader-Willi/Angelman (F), Langer-Giedion (G), Cri-du-chat (H), and Jacobsen (I) associated aMD. In panel A (20p13del), the PPV on FF of 1% is higher in 5 and 10M RPS compared to the 20M RPS due to the minimal number of true-positive (TP) MD calls (3 and 1) and zero false-positive (FP) MD calls compared to the 20M RPS where there was only one TP and one FP MD call (leading to a PPV of 0.5 for 20M RPS and 1 for 5 and 10M RPS).

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

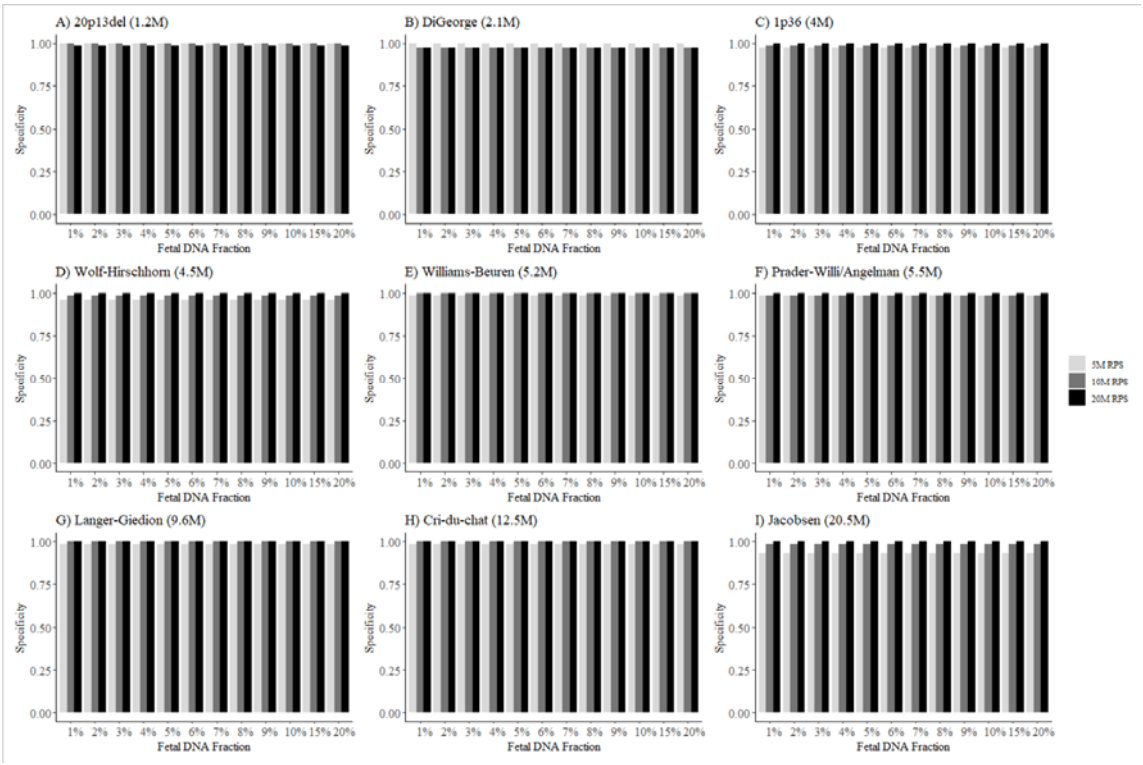
It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 3. BinDel inferred risk scores for euploid fetus samples on downsampled 5M reads per sample (RPS) and upsampled 20M RPS from 10M RPS. Red circles represent euploid samples with wrongly inferred MD high risk. The dashed line represents the cut-off point for high-risk calls in each graph. Data obtained with sequencing depths of 5M RPS (A) and 20M RPS (B) are shown.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

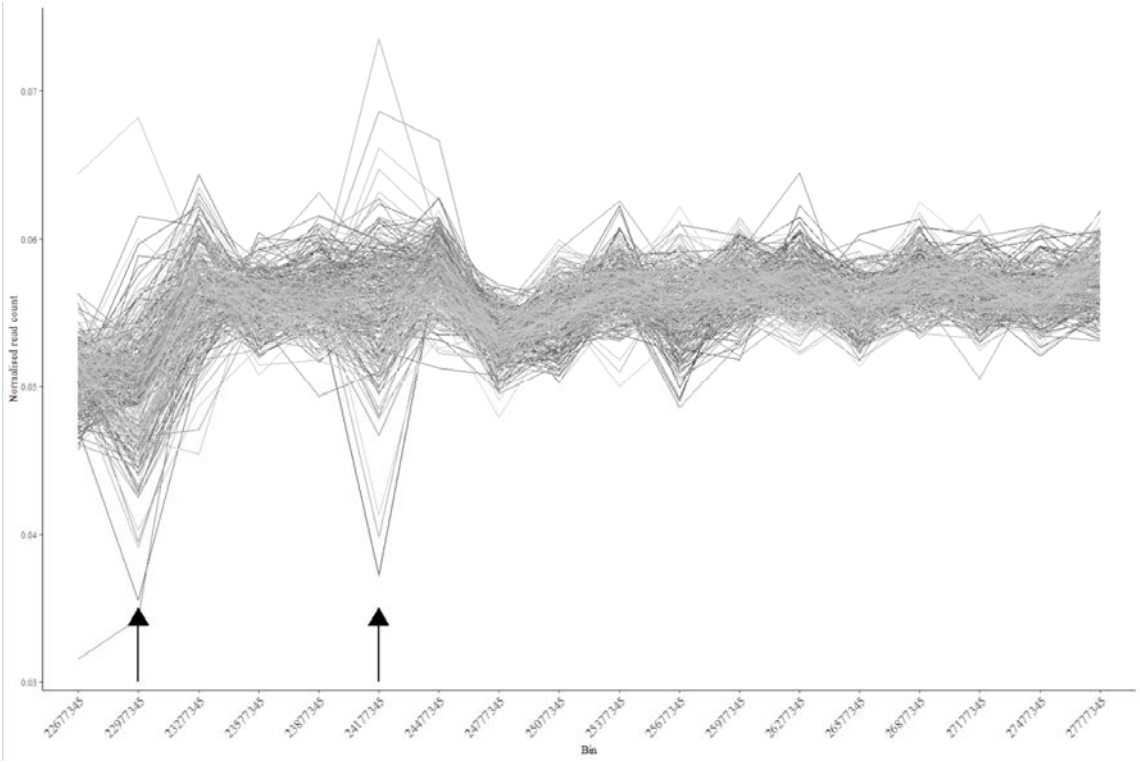
It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 4. BinDel specificity for in silico created microdeletion detection. Specificity estimates are grey, dark grey, and black for 5, 10 and 20M reads per sample (RPS). Specificity calculated for 20p13del (A), DiGeorge (B), 1p36 (C), Wolf-Hirschhorn (D), Williams-Beuren (E), Prader-Willi/Angelman (F), Langer-Giedion (G), Cri-du-chat (H), and Jacobsen (I) associated MD.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

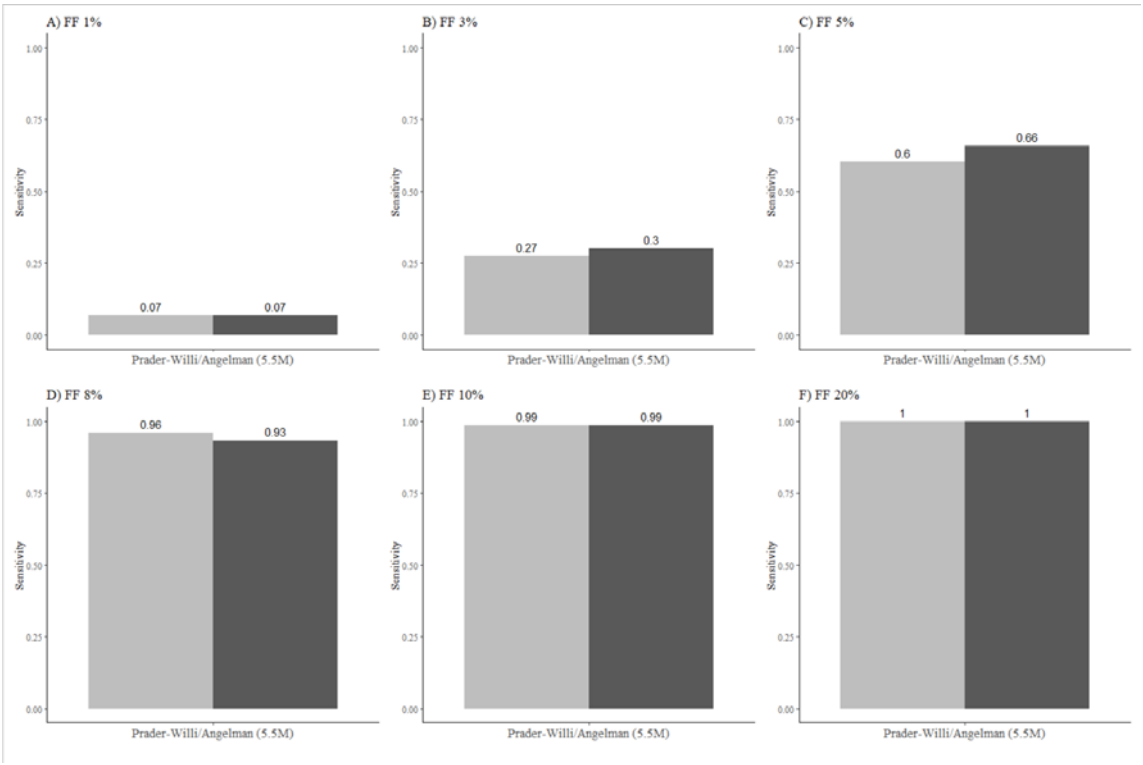
It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 5. 10M reads per sample (RPS) reference set (300 samples) bin normalised read count distribution in Prader-Willi/Angelman region. Each grey line represents a sample from the reference set. Each bin is 300 kb long and two highly variable bins 24177345 and 2297734 are indicated with black arrow.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

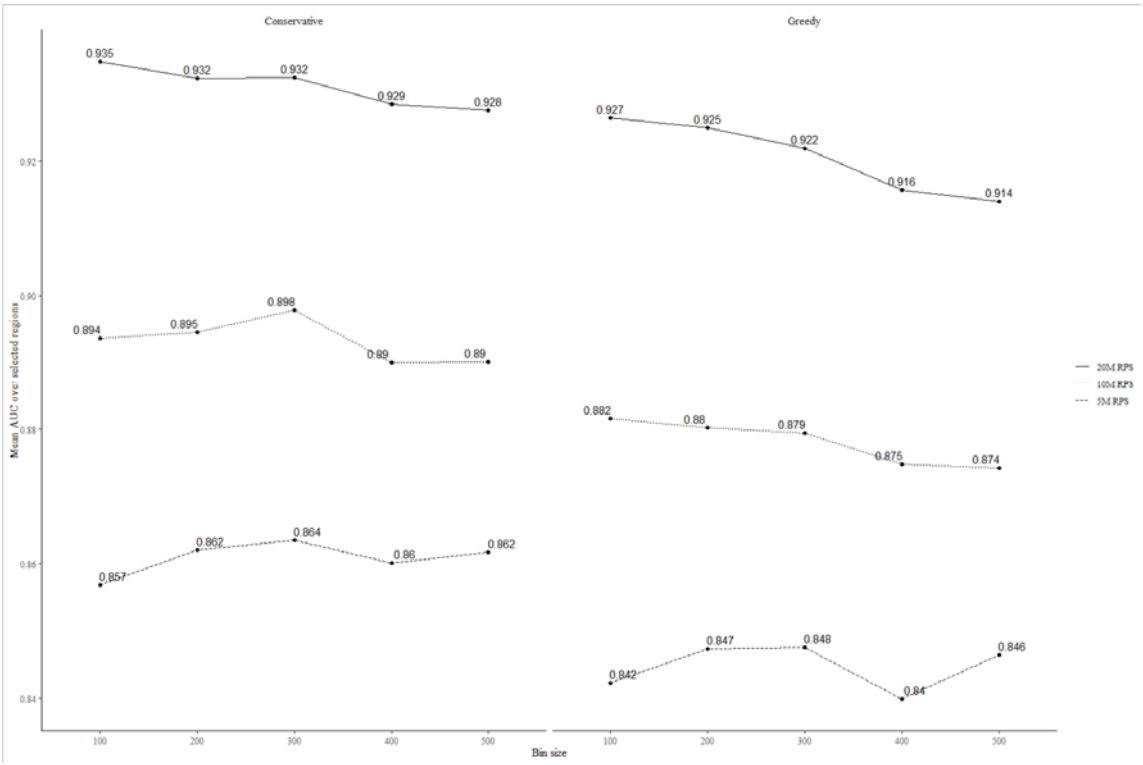
It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 6. BinDel sensitivity for detection of in silico created Prader-Willi/Angelman syndrome microdeletion with and without removal of bins 24177345 and 22977345 on 10M reads per sample (RPS). Estimates obtained at fetal fractions of 1% (A), 3% (B), 5% (C), 8% (D), 10% (E) and 20% (F). Dark grey sensitivity is without bins 24177345 and 22977345.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .



Suppl. Figure 7. Mean AUC of 9 MD regions on sequencing depths of 5, 10 and 20M RPS with bin sizes of 100-500 kb with greedy and conservative risk assessment over sample groups 2, 7, 8, 9, 10 and 11. 300 kb conservative bin size returned the highest sum of mean AUC over 5, 10 and 20M RPS and was selected as the bin size for the analysis.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Suppl. Table 1. Summary information of the analysed sample groups. The summary includes sample group identification ID, sample group size, initial sample group (if derived in silico from another sample group), how the sample group was used, average read count in millions of reads (M RPS), and sample group description.

Sample group	Size	Base sample group	Usage	Coverage	Description
1	300		Algorithm euploid reference set	10M RPS	Previously reported to be euploid by the NIPTIFY screening test and postnatal evaluation.
2	73		Algorithm development	10M RPS	See sample group 1.
3	244		Algorithm validation	20M RPS	Clinically validated samples from the previous study (except one sample with Prader-Willi syndrome) [12]. T13: 1 T18: 5 T21: 12 Prader-Willi: 1 (diagnosed postnatally) Euploid: 225
4	2		Microdeletion validation	20M RPS	SeraCare Life Sciences Inc ccfDNA-like mixture of human genomic DNA that consists of matched maternal and fetus DNA sequenced at 20M RPS. The sample group consists of samples SC005 (lot # 10446565) and SC0042 (lot # 10571706), which both consist of matched DNA of maternal and fetus with DiGeorge syndrome.
5	2		Microdeletion validation	10M RPS	SeraCare Life Sciences Inc ccfDNA-like mixture (see sample group 4). The sample group consists of two samples, SC005 (lot # 10446565) and SC016 (lot # 10560229), sequenced at 10M RPS. SC005 is a DiGeorge syndrome sample, and SC016 is a 20p13 deletion + 3q29 duplication sample.
6	2	5	Microdeletion validation	5M RPS	Subsampled sample group 5.
7	73	2	Algorithm	5M RPS	Subsampled sample group 2.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

			development		
8	73	2	Algorithm development	20M RPS	Upsampled from sample group 2 by combining initial samples randomly. More precisely, combinations of 73 samples were taken twice without repetition and merged.
9	730	7	Algorithm specificity, positive predictive value (PPV) and sensitivity analysis	5M RPS	Each microdeletion was bioinformatically created in silico to each base sample per each studied FF. FF: 1%, 3%, 5%, 8%, 10%, 20% In silico created microdeletions: 1. 1p36 2. Wolf-Hirschhorn 3. Cri-du-chat 4. Williams-Beuren 5. Langer-Giedion 6. Jacobsen 7. Angelman/Prader-Willi 8. 20p13 del 9. DiGeorge Given sample S, we transformed S into new samples of: $S_{1p36\ FF\ 1\%}$, $S_{1p36\ FF\ 3\%}$ ... $S_{DiGeorge\ FF\ 20\%}$
10	730	2	Algorithm specificity, PPV and sensitivity analysis	10M RPS	See sample group 9 description.
11	730	8	Algorithm specificity, PPV and sensitivity analysis	20M RPS	See sample group 9 description.

medRxiv preprint doi: https://doi.org/10.1101/2022.09.20.22280152; this version posted September 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Suppl. Table 2. Selected regions for analysis with available coordinates and coordinates used [18,19]**.**

Syndrome	Occurrence	DECIPHER (b38)	OMIM (b38)	SeraCare Life Sciences Inc SC106, SC005 specification (b37)	Coordinates used	Length
1p36 (deletion)	1:5000 [22]	10,001 12,780,116	23,600,000 27,600,000		OMIM	4Mb
3q29 (duplication)	NA	195,999,964 197,617,792	192,600,000 198,295,559	193,478,683 197,861,598	OMIM	5.7Mb
Wolf-Hirschhorn (deletion)	1:20,000-50,000 [23]	1,567,470 2,108,509	0 4,500,000		OMIM	4.5Mb
Cri-du-chat (deletion)	1:15,000-50,000 [3]	10,001 12,533,192			DECIPHER	12.5Mb
Williams-Beuren (deletion)	1:7,500 [24]	73,330,452 74,728,334	72,700,000 77,900,000		OMIM	5.2Mb
Langer-Giedion (deletion)	NA		116,700,000 126,300,000		OMIM	9.6Mb
Jacobsen (deletion)	1:100,000 [25]		114,600,000 135,086,622		OMIM	20.5Mb
Prader-Willi/ Angelman (deletion)	1:15,000–25,000 [26]	22,677,345/23,374,765 28,193,120			DECIPHER	5.5Mb
20p13 (deletion)	NA			60,747 1,292,456	SeraCare Life Sciences Inc	1.2Mb
DiGeorge (deletion)	1:1,524 [1]	19,022,279 21,098,156		18,894,339 21,440,514	DECIPHER	2.1Mb