

RESEARCH ARTICLE

High Levels of Sample-to-Sample Variation Confound Data Analysis for Non-Invasive Prenatal Screening of Fetal Microdeletions

Tianjiao Chu^{1,2}, Suveyda Yeniterzi^{1,2}, Svetlana A. Yatsenko^{1,2}, Mary Dunkel², Patricia A. Shaw², Kimberly D. Bunce², David G. Peters^{1,2*}

1 Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, Pittsburgh, PA, United States of America, **2** Center for Fetal Medicine, Magee-Womens Research Institute, Pittsburgh, PA, United States of America

* dgp6@pitt.edu



OPEN ACCESS

Citation: Chu T, Yeniterzi S, Yatsenko SA, Dunkel M, Shaw PA, Bunce KD, et al. (2016) High Levels of Sample-to-Sample Variation Confound Data Analysis for Non-Invasive Prenatal Screening of Fetal Microdeletions. PLoS ONE 11(6): e0153182. doi:10.1371/journal.pone.0153182

Editor: Maj Hulten, Warwick University, UNITED KINGDOM

Received: February 18, 2016

Accepted: March 24, 2016

Published: June 1, 2016

Copyright: © 2016 Chu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by grants from the Magee-Womens Research Institute and Foundation (MWRIF) and the NIH (R01 HD068578) (both to DGP).

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Our goal was to test the hypothesis that inter-individual genomic copy number variation in control samples is a confounding factor in the non-invasive prenatal detection of fetal microdeletions via the sequence-based analysis of maternal plasma DNA. The database of genomic variants (DGV) was used to determine the “Genomic Variants Frequency” (GVF) for each 50kb region in the human genome. Whole genome sequencing of fifteen karyotypically normal maternal plasma and six CVS DNA controls samples was performed. The coefficient of variation of relative read counts (cv.RTC) for these samples was determined for each 50kb region. Maternal plasma from two pregnancies affected with a chromosome 5p microdeletion was also sequenced, and analyzed using the GCREM algorithm. We found strong correlation between high variance in read counts and GVF amongst controls. Consequently we were unable to confirm the presence of the microdeletion via sequencing of maternal plasma samples obtained from two sequential affected pregnancies. Caution should be exercised when performing NIPT for microdeletions. It is vital to develop our understanding of the factors that impact the sensitivity and specificity of these approaches. In particular, benign copy number variation amongst controls is a major confounder, and their effects should be corrected bioinformatically.

Introduction

Definitive prenatal diagnosis of genetic disease is currently performed via amniocentesis (AF) or chorionic villus sampling (CVS). These are invasive procedures that carry an inherent risk of miscarriage, fetal morbidity and parental anxiety[1–7]. Because of these risks, there has been considerable interest in the development of non-invasive tests that can be performed on a maternal blood sample in early gestation. One promising approach involves the analysis of cell-free fetal DNA in maternal plasma, which can serve as a substrate for fetal genotyping[8]. These nucleic acids are thought to arise from apoptotic cells known as syncytiotrophoblastic

microparticles that are shed from the placental villus into the uterine circulation[9–12]. Progress in this field has rapidly accelerated following the demonstration that whole genome sequencing of maternal plasma DNA can be used for the non-invasive prenatal testing (NIPT) of common aneuploidy, including successful translation of these methods into the clinical setting[13–15].

Other chromosomal abnormalities such as microdeletions and microduplications occur with approximately twice the frequency of aneuploidy[16, 17] and are, therefore, of major clinical significance. Because of this, a number of methods for the diagnosis of microdeletions and microduplications have been developed. However, the majority of these anomalies are smaller in size than the limits of detection of metaphase chromosome analysis. Thus, array comparative genomic hybridization (aCGH) has emerged as a powerful tool for their high-resolution detection. Unfortunately, however, aCGH must be performed utilizing AF or CVS.

In 2011 we demonstrated proof of concept that NIPT can be used for the non-invasive detection of fetal microdeletions[18] and this has since been confirmed by at least three other groups[19–21]. NIPT for microdeletions and microduplications is technically challenging, with complications arising from the potential for false positive and false negative results, unanticipated findings, and relatively high cost due to the amount of sequence data required. Furthermore, our understanding of the utility of this approach in a broad sense is limited. In light of this we tested the hypothesis that inter-individual variation in genomic copy number of control DNA samples is a confounding factor in the non-invasive prenatal detection of fetal microdeletions via the sequence-based analysis of maternal plasma DNA. We furthermore describe the analysis of a 318.46kb interstitial chromosome 5p microdeletion and highlight a number of key observations relating to NIPT in this context.

Materials and Methods

Human tissue samples

All patients provided written informed consent for use of their samples and clinical data. These women received prenatal care at Magee-Womens Hospital within the years 2008–2013, and were recruited from clinics and private practices. This protocol was approved by the University of Pittsburgh Institutional Review Board (Protocol # IRB PRO07090033). A family presented to Magee-Womens Hospital of UPMC at the Medical Genetics Department. The mother was Gravida eight, Para three, and was seen at 19.0 weeks gestational age due to an increased risk for Down Syndrome (1:42) and open neural tube defect (1:160) by multiple marker screen. An amniocentesis was performed, and during ultrasound for the procedure, bilateral club feet and an echogenic bowel were noted. Karyotype and FISH analysis from the amniocentesis revealed 46,XX. Comparative genomic hybridization (CGH)-based microarray analysis revealed a 318.46kb loss of 5p15.33 (Fig 1).

DNA extraction and sequencing

Plasma handling, DNA extraction and DNA sequencing were performed as previously described[18].

Data analysis

Data were analyzed using the previously described GCREM approach[22]. Briefly, GCREM is a linear mixed effect model for the read counts of different genomic regions in a DNA library, where the GC content is an independent variable with a library specific random coefficient. This linear mixed effect model is fitted using a set of libraries with known normal genomes to

Figure 1A

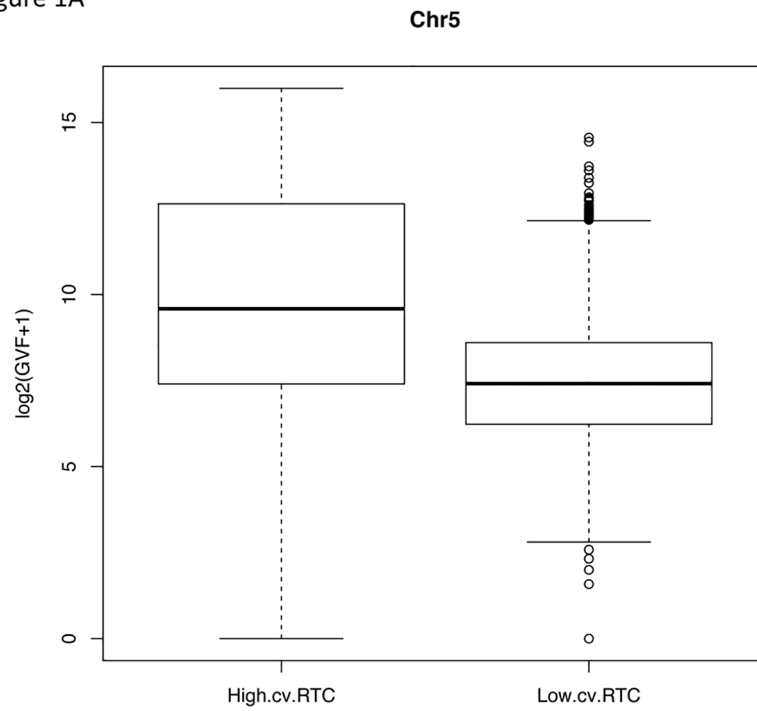


Figure 1B

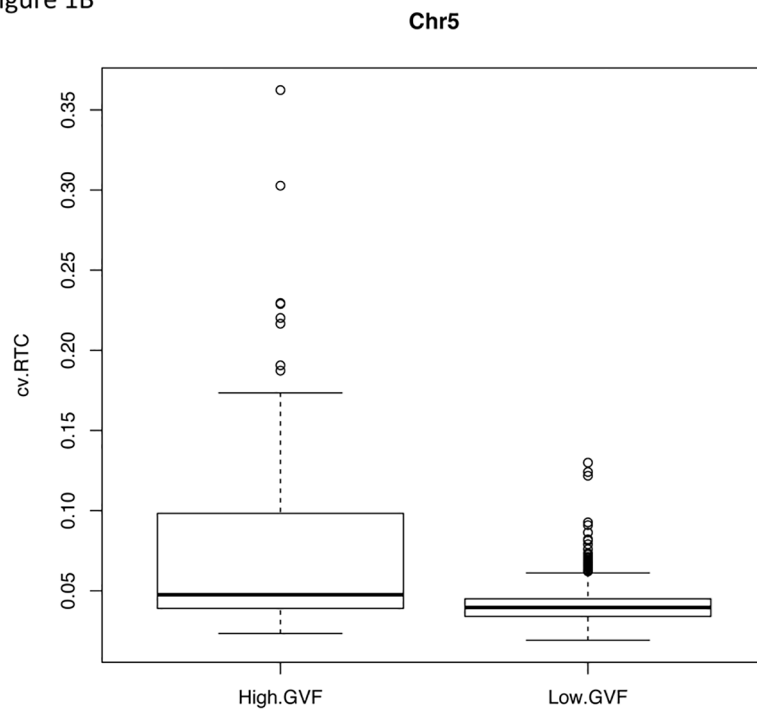


Fig 1. Comparative Genomic Hybridization Analysis of Affected Fetal Sample (First Pregnancy). Array CGH profile showing an interstitial deletion in the short arm of chromosome 5. Top: Ideogram of chromosome 5. The deleted 5p15.33 region is indicated by a red rectangle. Below: A magnified view of the 5p subtelomeric

region. Positions are displayed according to GRCh37/hg19 Genome Browser. Shaded blue area indicates a loss in DNA copy number detected by 22 oligonucleotide probes (blue dots), located in the interval chr5:174,979–493,441 and encompassing an approximately 318 kb segment.

doi:10.1371/journal.pone.0153182.g001

estimate the region specific corrections for a list of genomic regions of interest. Then we apply the corrections to the target library, and fit a linear model using only the regions that are believed to be normal in the target library. We then predict the read count for the regions of interest. By comparing the observed read counts with the predicted read counts, we can test if the suspected regions indeed carry insertions or deletions.

Results

Summary statistics for sequencing libraries described herein are shown in [S1 Table](#).

High variance in sequencing read counts amongst normal samples strongly correlates with the frequency of genomic copy number variants

We hypothesized that the locus-specific variation of sequence read counts in libraries from clinically normal samples caused by benign copy number variation will negatively affect the performance of sequencing based copy number alteration (CNA) detection algorithms. Common to all sequencing based CNA detection algorithms, the test for the presence of a CNA is dependent on three parameters: T: a measurement of the normalized read count of the region of interest in the target sample; C: the average of the normalized read count in that region over a group of normal control samples; S: the standard deviation of the normalized read count for that group of normal control samples. The test statistic usually can be expressed as $(T-C)/S$. With benign copy number variants in the group of normal control samples, we could see a biased C and an inflated S, both of which will reduce the sensitivity and specificity of the CNA detection algorithm.

To examine the relation between the variation of sequence read counts in libraries from clinically normal samples and benign copy number variation, we selected fifteen maternal plasma libraries from normal pregnancy samples that had been corrected for GC bias ([S1 Fig](#)). We divided the human reference genome (GRCh37) into 50kb non-overlapping regions, and calculated the relative read count (RTC) for each 50kb region, which is defined as the ratio of the read count in a given 50kb region to the median of the read counts in all 50kb regions in the whole genome (excluding regions with no aligned reads in all libraries). We also downloaded the Database for Genomic Variants (DGV) [[23](#)]. We first counted, for each 5kb region in human genome, the number of structural variants reported in the DGV database overlapping with that 5kb region (NDGV). Then we calculated, for each 50kb region, the Genomic Variants Frequency (GVF) as the sum of NDGV over the ten 5kb regions inside the 50kb region. Recall that GVF is the number of genomic variants observed in each 50kb region. It is \log_2 transformed so that its distribution is closer to a Gaussian distribution. We found a strong dependency between the coefficient of variation of the RTC ($cv.RTC$) across the fifteen libraries for each 50kb region, and the risk of structural variation, measured as $\log_2(GVF+1)$, for that 50kb region. For example, in each autosome, regions with high $cv.RTC$ ($\geq 95\%$ quantile of the $cv.RTC$ of all 50kb regions in that chromosome) have GVF values that are significantly higher than regions with low $cv.RTC$ ($< 95\%$ quantile) (Wilcoxon two sample test) (Figs [2A](#) and [3B](#), [S2 Fig](#) and [S2A Table](#)). On the other hand, the $cv.RTC$ in regions with high GVF [$\log_2(GVF+1) \geq 95\%$ quantile of $\log_2(GVF+1)$ of all 50kb regions in each chromosome] is

Figure 2A

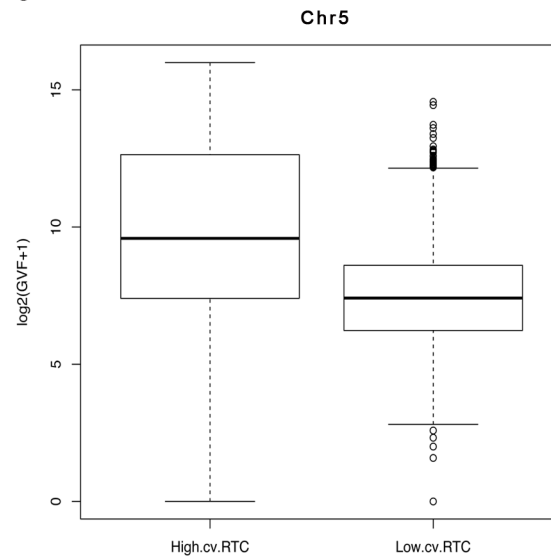


Figure 2B

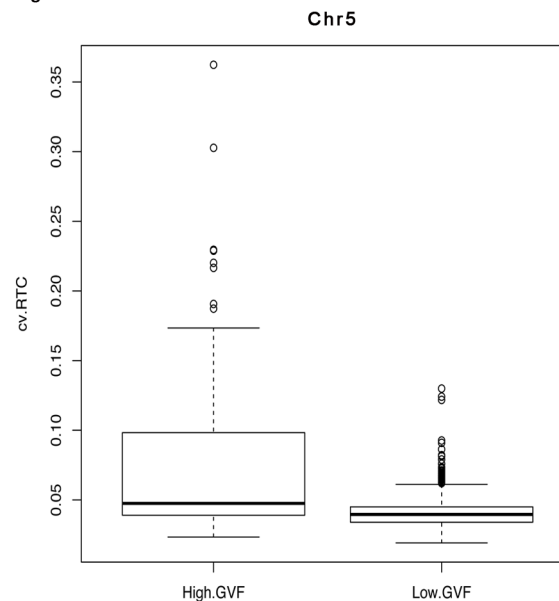


Fig 2. Boxplots of Genomic Variants Frequency (GVF) and coefficient of variation of Relative Read Counts (cv.RTC). (A) Boxplots of $\log_2(\text{GVF}+1)$ for regions with high cv.RTC (above 95% quantile of the cv.RTC of all regions on chromosome 5) and low cv.RTC (below 95% quantile of cv.RTC) in maternal plasma control libraries on chromosome 5 (representative). (B) Boxplot of cv.RTC of the maternal plasma control libraries for regions with high GVF (above 95% quantile of the GVF of all regions on chromosome 5) and low GVF (below 95% quantile of GVF) on chromosome 5 (representative).

doi:10.1371/journal.pone.0153182.g002

significantly higher than in regions with low GVF [$\log_2(\text{GVF}+1) < 95\%$ quantile] (Figs 2B and 3B, S2 Fig and S3A Table).

The high sample-to-sample variation observed in the fifteen plasma samples in certain regions is not coincidental. We further analyzed the six unrelated control CVS obtained between 10 and 13 weeks gestation, and found high sample-to-sample variation in the same regions. More precisely, after correction for GC bias, for each of the 22 autosomes, we found

Figure 3A

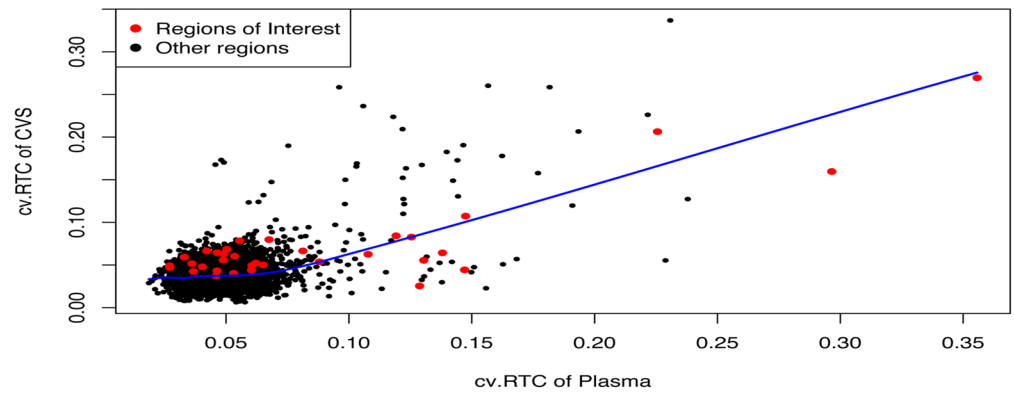


Figure 3B

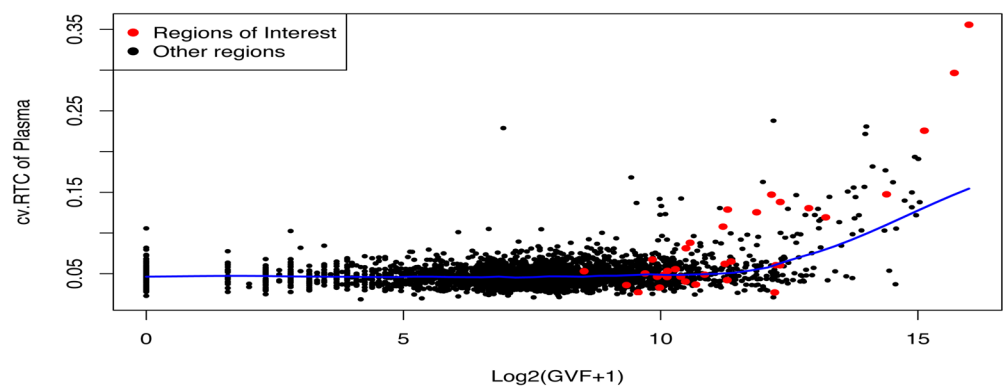


Figure 3C

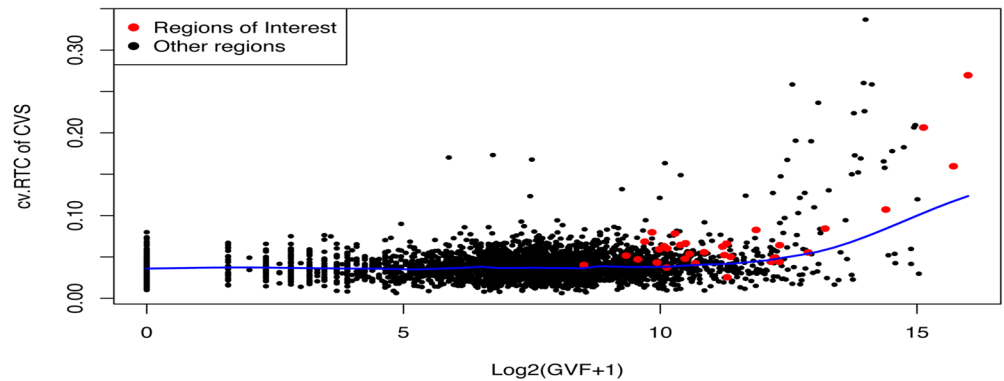


Fig 3. Plots of Coefficient of Variation of the Relative Read Count (cv.RTC) for Chromosome 5. (A) The cv.RTC of the 50kb regions in the CVS control libraries (y axis) against that in the maternal plasma control libraries (x axis) for chromosome 5. (B) The cv.RTC of the 50kb regions in the plasma control libraries (y axis) against the $\log_2(\text{GVF}+1)$ of the 50kb regions reported in the DGV database (x axis) for chromosome 5. (C) The cv.RTC of the 50kb regions in the CVS control libraries (y axis) against the $\log_2(\text{GVF}+1)$ of the 50kb regions reported in the DGV database (x axis) for chromosome 5. In all three plots, red dots represent 50kb regions around the microdeletion (see Fig 3), black dots represent other 50kb regions in chromosome 5.

doi:10.1371/journal.pone.0153182.g003

that the cv.RTC of the 50kb regions in the CVS samples is significantly and positively correlated with that in the plasma samples (t test, p values < 2.2×10^{-16}) (Fig 3A, S2 Fig), and shows a strong dependency with the risk of structural variation ($\log_2(\text{GVF}+1)$) in the corresponding regions (Fig 3C, S2 Fig, S2B and S3B Tables). We also explored other data preprocessing

Table 1. The 25%, 50%, 75%, and 97.5% quantiles of the coefficient of variation of RTC in chromosome 5. For CVS, the coefficient of variations were estimated from 6 libraries. For maternal plasma, the coefficient of variations were estimated from fifteen libraries unrelated to the CVS samples.

Sample	GC correction	Allow mismatch	25.00%	50.00%	75.00%	97.50%
CVS	Yes	No	0.0286	0.0370	0.0471	0.0784
CVS	Yes	Yes	0.0231	0.0297	0.0369	0.0632
CVS	No	No	0.0293	0.0381	0.0487	0.0871
CVS	No	Yes	0.0239	0.0307	0.0386	0.0738
Plasma	Yes	No	0.0383	0.0464	0.0563	0.1131
Plasma	Yes	Yes	0.0324	0.0391	0.0475	0.1027
Plasma	No	No	0.0463	0.0554	0.0676	0.1214
Plasma	No	Yes	0.0402	0.0481	0.0593	0.1151

doi:10.1371/journal.pone.0153182.t001

options, including deriving RTC from the read count data without GC content correction, as well as using different parameters for the alignment, e.g., allowing mismatch vs not allowing mismatch, and found similar results (Table 1). Interestingly, the CVS samples tend to have smaller coefficients of variation than the plasma samples. Allowing mismatch and correcting for GC content can also reduce the coefficient of variation, especially for the plasma samples (Table 1).

Note that the samples included in the DGV database are not necessarily representative of the population from which we collected the plasma and CVS control samples. This makes the strong correlation between the cv.RTC of the plasma and CVS control libraries on the one side, and the GVF derived from the DGV database on the other side even more remarkable.

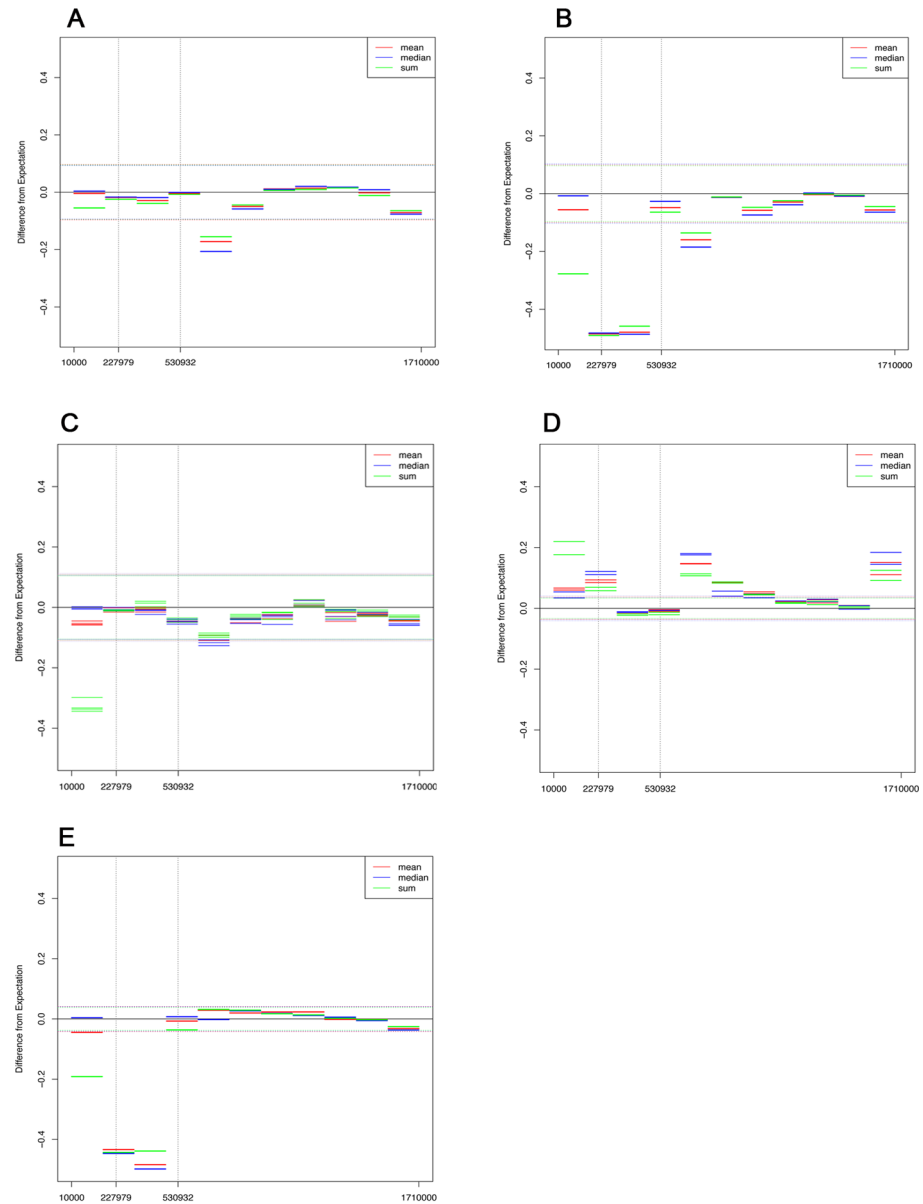
The effect of high variance of read counts on the ability to detect a fetal microdeletion by sequencing maternal plasma DNA

The effect of high variance of the read counts on the non-invasive prenatal diagnosis of fetal microdeletion is best illustrated by the following case study. A plasma DNA sample (PL741) from a mother carrying a fetus with a confirmed paternally inherited ~318.46Kb interstitial deletion on chromosome 5 between 174,979–493,441 (hg19) (Fig 1) was obtained at 32.4 weeks gestation and subjected to paired end (100 x 2) whole genome sequencing on the Illumina HiSeq 2000.

We first tested for the presence of the microdeletion in PL741 using the fifteen confirmed karyotypically normal control plasma samples discussed above. As shown in Fig 4A, there was no clear evidence that the PL741 microdeletion was detectable using this approach.

We also sequenced DNA from fetal amniocytes (C741), maternal leukocytes (M741), and a paternal sample (Pat741) obtained from the affected PL741 pregnancy. We used the six karyotypically normal CVS tissue samples discussed above as controls, so that the pure maternal, paternal, and fetal samples from pregnancy PL741 could be compared with “like”, high molecular weight nuclear DNA, controls rather than plasma-derived DNA. As shown in Fig 4B, we were clearly able to detect copy number loss in the confirmed 318.46Kb deleted region in the pure fetal (C741). As expected, consistent with paternal transmission of this mutation, we found in the paternal sample (Pat741) an almost identical pattern of copy number loss (Fig 4E) to the pure fetal sample, but no evidence of copy number loss in the maternal leukocyte sample (M741) (Fig 4C).

In order to confirm that fetal DNA in the PL741 plasma was indeed at a detectable level and within the expected range of sensitivity of our assay we analyzed the sequencing data from the PL741 plasma sample to determine the percentage of fetal genome equivalents. We used a



GCREM test results.

- (A) Maternal Plasma DNA Sample from the First Pregnancy (PL741). (B) Fetal DNA Sample from the First Pregnancy (C741) (C) Maternal Genomic DNA (Leukocyte-Derived) Sample (M741) (D) Maternal Plasma DNA Sample from the Second Pregnancy (PL1694) (E) Paternal DNA Sample (pat741)

Fig 4. GCREM Tests of Chromosome 5 Microdeletion. (A) GCREM test of the maternal plasma library (data is from 1 of a total of 13 HiSeq 2000 lanes) from the first pregnancy. (B) GCREM test of the fetal amniocyte library from the first pregnancy. (C) GCREM test of the maternal genomic libraries (data is from 4 HiSeq 2000 lanes) from the first pregnancy. (D) GCREM test of the maternal plasma library from the second pregnancy (PL1694). (E) GCREM test of the paternal genomic (leukocyte-derived) library (pat741). For all plots, Y axis is the difference between observed RTC and expected RTC. Positive values mean potential gain, negative values mean potential loss. X axis is the position on the chromosome. Each horizontal bar represents a 150kb region. Dotted lines are standard deviation of the expected RTC.

doi:10.1371/journal.pone.0153182.g004

method that we developed for this purpose, which is a modification of our previously published work[24]. Briefly, we identified SNPs that are homozygous in the maternal sample and heterozygous in the fetal sample. The frequency of the paternally inherited allele to the maternal allele then is half of the fetal frequency. Using generalized linear model, our estimation of the fetal DNA frequency in the first plasma data is approximately 15%, which is well within the expected normal range for such samples.

Eventually, we were able to obtain a chorionic villus sample from a subsequent pregnancy involving the same family. Again, microarray analysis was performed (not shown) and the same chromosome 5 microdeletion was identified. A plasma sample (PL1694) was obtained at 11 weeks. GCREM analysis of the maternal plasma (PL1694) from the second pregnancy showed the same results as those (PL741) from the first pregnancy in that the 318.46Kb deletion was not detected (Fig 4D).

A close examination of the fifteen plasma and six CVS control samples showed that, despite corrected for GC bias (S1 Fig), the read counts of the regions of interest in these controls displayed extremely high sample-to-sample variation. We found that, across the control samples, the coefficient of variation of the RTC in the regions of interest are much higher than other regions of chromosome 5. For example, for the RTC derived from the GC corrected read count data, the lower quartile, median, upper quartile, and 97.5% of the coefficient of variation for all chromosome 5 regions are 0.038, 0.046, 0.056, and 0.113 respectively. However, among the 31 regions of interest (located in chr5:10000–1710000), 17 of them have a coefficient of variation greater than the upper quartile, and 11 of them have a coefficient of variation greater than the 97.5% quantile.

Discussion

We present an analysis of sequenced normal maternal plasma and CVS DNA samples that were used as controls in the context of NIPT for fetal microdeletions and microduplications. We found that in some genomic regions, these control samples exhibited extremely high variance, as observed in both the CVS and plasma control data. The high variance in these regions is not incidental, because the distribution of high variance regions across the whole genome is strikingly similar between the fifteen maternal plasma libraries and the six unrelated CVS libraries. More importantly, we found a strong correlation between the distribution of the high variance regions in both the plasma and CVS libraries and the distribution of reported structural variation in the DGV database. Statistical tests showed that the RTCs of the regions where the DGV database reported a high degree of structural variation have significantly higher coefficient of variation than the regions with fewer reported structural variations.

The presence of regions with high risk of structural variations in the control libraries represents a serious challenge to the application of NIPT technologies for the detection of microdeletions and microduplications. While the number of high-risk regions may be small, the chance that a structural variation will occur in such regions is high. Using the data from the DGV database, about 7% of the 50kb regions were found to have a GVF of 2000 or higher, but these account for 70% of total GVF reported in the database.

To illustrate, we used those fifteen plasma samples as controls to investigate a maternal plasma sample with microdeletion on the short arm of chromosome 5. Because of the extremely high variance in the control samples over this region, we were unable to detect this via non-invasive methods, despite the comprehensive analysis of plasma DNA using proven statistical tools.

Based on the results presented above, in order to apply NIPT technology to the detection of a broad spectrum of microdeletions and microduplications, the influence of structural

variations in the control plasma samples would need to be mitigated. To do this, we would ideally need to obtain a subset of plasma samples with no structural variants, at least over the regions of interest, so that these samples could serve as the controls for the microdeletion and microduplication tests. However, given the diversity of the human genome, it is difficult, if not impossible, to find samples containing sequence over a large region that is identical to the reference genome. Instead, we recommend correcting the effect of structural variants on a given maternal plasma-sequencing library such that a corrected library will have a read count distribution over the regions of interest similar to that of a sample of the reference genome.

Briefly, this would require matched maternal plasma (MP), maternal blood cells (MBC), and CVS samples. By deep sequencing of the maternal and fetal genomes, preferably enriched over the regions of interest, we can assemble the maternal and fetal genomes against the reference genome. Importantly, because the mother and the fetus share a copy of chromosome sets, we can derive the haplotype information for both the mother and the fetus over the regions of interest. By sequencing the maternal plasma samples, we can also estimate the percentage of fetal DNA in maternal plasma[24]. Then we can determine either analytically or by simulation how the structural variants in the maternal and fetal genomes affect, in maternal plasma sequencing libraries, the distribution of reads over the reference genome. In turn we can correct the maternal plasma libraries to remove these effects so that the read distribution in the corrected libraries would be similar to that from libraries generated from the reference genome. The availability of control plasma samples with corrected read counts over the regions of interest will greatly increase the range of application of the NIPT technology from regions with low risk of copy number variation to regions of higher risk of copy number variation, so that it can be offered as a viable and risk-free alternative to invasive diagnostic procedures.

Supporting Information

S1 Fig. GC Correction Curves for Fifteen Karyotypically Normal Control Plasma Samples and 6 Karyotypically Normal CVS Samples. The curves represent a nonparametric estimation of the relationship between the read count in a genomic region and the GC content of that genomic region. The X axis is the GC content of 50kb genomic regions. The Y axis is the number of reads aligned to a 50kb region.

(PDF)

S2 Fig. Correlation Between Read Count Variance and Frequency of Structural Variation for All Chromosomes. There are 22 groups of plots in this file for the 22 autosomes. Each group has three plots. In each group, the first plot is the cv.RTC of the 50kb regions in the CVS control libraries against that in the maternal plasma control libraries for each autosome. The second plot is the cv.RTC of the 50kb regions in the plasma control libraries against the log₂ (GVF+1) of the 50kb regions reported in the DGV database. The third plot is the cv.RTC of the CVS control libraries against the DGV database.

(PDF)

S1 Table. Summary statistics for Sequencing Libraries.

(XLS)

S2 Table. (A). Comparison of the number of structural variants reported in the DGV database between regions in fifteen maternal plasma control libraries displaying high and low cv.RTC
(B). Comparison of the number of structural variants reported in the DGV database between regions in six CVS controls libraries displaying high and low cv.RTC.

(DOCX)

S3 Table. (A). Comparison of the coefficient of variation of relative read count between regions in fifteen plasma controls libraries displaying high and low numbers of reported structural variants in the DGV database. (B). Comparison of the coefficient of variation of relative read count between regions in six CVS controls libraries displaying high and low numbers of reported structural variants in the DGV database. (DOCX)

Author Contributions

Conceived and designed the experiments: TC DGP. Performed the experiments: TC DGP PAS KDB. Analyzed the data: TC SY SAY DGP. Contributed reagents/materials/analysis tools: TC DGP. Wrote the paper: TC DGP. Acquisition of data: MD.

References

1. Mujezinovic F, Alfirevic Z. Procedure-related complications of amniocentesis and chorionic villous sampling: a systematic review. *Obstet Gynecol*. 2007; 110(3):687–94. Epub 2007/09/04. doi: 110/3/687 [pii] doi: [10.1097/01.AOG.0000278820.54029.e3](https://doi.org/10.1097/01.AOG.0000278820.54029.e3) PMID: [17766619](https://pubmed.ncbi.nlm.nih.gov/17766619/).
2. Tabor A, Alfirevic Z. Update on procedure-related risks for prenatal diagnosis techniques. *Fetal diagnosis and therapy*. 27(1):1–7. Epub 2010/01/07. doi: 000271995 [pii] doi: [10.1159/000271995](https://doi.org/10.1159/000271995) PMID: [20051662](https://pubmed.ncbi.nlm.nih.gov/20051662/).
3. Tabor A, Philip J, Madsen M, Bang J, Obel EB, Norgaard-Pedersen B. Randomised controlled trial of genetic amniocentesis in 4606 low-risk women. *Lancet*. 1986; 1(8493):1287–93. Epub 1986/06/07. PMID: [2423826](https://pubmed.ncbi.nlm.nih.gov/2423826/).
4. Odibo AO, Gray DL, Dicke JM, Stamilio DM, Macones GA, Crane JP. Revisiting the fetal loss rate after second-trimester genetic amniocentesis: a single center's 16-year experience. *Obstet Gynecol*. 2008; 111(3):589–95. Epub 2008/03/04. doi: 111/3/589 [pii] doi: [10.1097/AOG.0b013e318162eb53](https://doi.org/10.1097/AOG.0b013e318162eb53) PMID: [18310360](https://pubmed.ncbi.nlm.nih.gov/18310360/).
5. Hewison J, Nixon J, Fountain J, Hawkins K, Jones CR, Mason G, et al. A randomised trial of two methods of issuing prenatal test results: the ARIA (Amniocentesis Results: Investigation of Anxiety) trial. *BJOG*. 2007; 114(4):462–8. Epub 2007/03/24. doi: BJO1276 [pii] doi: [10.1111/j.1471-0528.2007.01276.x](https://doi.org/10.1111/j.1471-0528.2007.01276.x) PMID: [17378819](https://pubmed.ncbi.nlm.nih.gov/17378819/).
6. Hewison J, Nixon J, Fountain J, Cocks K, Jones C, Mason G, et al. Amniocentesis results: investigation of anxiety. The ARIA trial. *Health Technol Assess*. 2006; 10(50):iii, ix-x, 1–78. Epub 2006/12/01. doi: 99/48/04 [pii]. PMID: [17134598](https://pubmed.ncbi.nlm.nih.gov/17134598/).
7. Hertling-Schaal E, Perrotin F, de Poncheville L, Lansac J, Body G. [Maternal anxiety induced by prenatal diagnostic techniques: detection and management]. *Gynecol Obstet Fertil*. 2001; 29(6):440–6. Epub 2001/07/21. doi: S1297958901001588 [pii]. PMID: [11462960](https://pubmed.ncbi.nlm.nih.gov/11462960/).
8. Lo YM, Tsui NB, Chiu RW, Lau TK, Leung TN, Heung MM, et al. Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. *Nat Med*. 2007; 13(2):218–23. Epub 2007/01/09. doi: nm1530 [pii] doi: [10.1038/nm1530](https://doi.org/10.1038/nm1530) PMID: [17206148](https://pubmed.ncbi.nlm.nih.gov/17206148/).
9. Chim SS, Tong YK, Chiu RW, Lau TK, Leung TN, Chan LY, et al. Detection of the placental epigenetic signature of the maspin gene in maternal plasma. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(41):14753–8. Epub 2005/10/06. doi: 0503335102 [pii] doi: [10.1073/pnas.0503335102](https://doi.org/10.1073/pnas.0503335102) PMID: [16203989](https://pubmed.ncbi.nlm.nih.gov/16203989/); PubMed Central PMCID: PMC1253547.
10. Orozco AF, Bischoff FZ, Horne C, Popek E, Simpson JL, Lewis DE. Hypoxia-induced membrane-bound apoptotic DNA particles: potential mechanism of fetal DNA in maternal plasma. *Ann N Y Acad Sci*. 2006; 1075:57–62. Epub 2006/11/17. doi: 1075/1/57 [pii] doi: [10.1196/annals.1368.007](https://doi.org/10.1196/annals.1368.007) PMID: [17108192](https://pubmed.ncbi.nlm.nih.gov/17108192/).
11. Orozco AF, Jorgez CJ, Horne C, Marquez-Do DA, Chapman MR, Rodgers JR, et al. Membrane-protected apoptotic trophoblast microparticles contain nucleic acids: relevance to preeclampsia. *Am J Pathol*. 2008; 173(6):1595–608. Epub 2008/11/01. doi: S0002-9440(10)61545-3 [pii] doi: [10.2353/ajpath.2008.080414](https://doi.org/10.2353/ajpath.2008.080414) PMID: [18974299](https://pubmed.ncbi.nlm.nih.gov/18974299/); PubMed Central PMCID: PMC2626372.
12. Gupta AK, Holzgreve W, Huppertz B, Malek A, Schneider H, Hahn S. Detection of fetal DNA and RNA in placenta-derived syncytiotrophoblast microparticles generated in vitro. *Clinical chemistry*. 2004; 50(11):2187–90. Epub 2004/10/27. doi: 50/11/2187 [pii] doi: [10.1373/clinchem.2004.040196](https://doi.org/10.1373/clinchem.2004.040196) PMID: [15502097](https://pubmed.ncbi.nlm.nih.gov/15502097/).

13. Palomaki GE, Deciu C, Kloza EM, Lambert-Messerlian GM, Haddow JE, Neveux LM, et al. DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: an international collaborative study. *Genetics in medicine: official journal of the American College of Medical Genetics*. 2012; 14(3):296–305. Epub 2012/01/28. doi: [10.1038/gim.2011.73](https://doi.org/10.1038/gim.2011.73) PMID: [22281937](https://pubmed.ncbi.nlm.nih.gov/22281937/).
14. Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE, Neveux LM, Ehrich M, et al. DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. *Genetics in medicine: official journal of the American College of Medical Genetics*. 2011; 13(11):913–20. Epub 2011/10/19. doi: [10.1097/GIM.0b013e3182368a0e](https://doi.org/10.1097/GIM.0b013e3182368a0e) PMID: [22005709](https://pubmed.ncbi.nlm.nih.gov/22005709/).
15. Ehrich M, Deciu C, Zwiefelhofer T, Tynan JA, Cagasan L, Tim R, et al. Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *American journal of obstetrics and gynecology*. 2011; 204(3):205 e1–11. Epub 2011/02/12. doi: [S0002-9378\(11\)00018-4](https://doi.org/S0002-9378(11)00018-4) [pii] doi: [10.1016/j.ajog.2010.12.060](https://doi.org/10.1016/j.ajog.2010.12.060) PMID: [21310373](https://pubmed.ncbi.nlm.nih.gov/21310373/).
16. Hillman SC, Pretlove S, Coomarasamy A, McMullan DJ, Davison EV, Maher ER, et al. Additional information from array comparative genomic hybridization technology over conventional karyotyping in prenatal diagnosis: a systematic review and meta-analysis. *Ultrasound Obstet Gynecol*. 2011; 37(1):6–14. Epub 2010/07/27. doi: [10.1002/uog.7754](https://doi.org/10.1002/uog.7754) PMID: [20658510](https://pubmed.ncbi.nlm.nih.gov/20658510/).
17. Shaffer LG, Dabell MP, Fisher AJ, Coppinger J, Bandholz AM, Ellison JW, et al. Experience with microarray-based comparative genomic hybridization for prenatal diagnosis in over 5000 pregnancies. *Prenatal diagnosis*. 2012; 32(10):976–85. Epub 2012/08/07. doi: [10.1002/pd.3945](https://doi.org/10.1002/pd.3945) PMID: [22865506](https://pubmed.ncbi.nlm.nih.gov/22865506/); PubMed Central PMCID: [PMC3491694](https://pubmed.ncbi.nlm.nih.gov/PMC3491694/).
18. Peters D, Chu T, Yatsenko SA, Hendrix N, Hogge WA, Surti U, et al. Noninvasive prenatal diagnosis of a fetal microdeletion syndrome. *The New England journal of medicine*. 2011; 365(19):1847–8. Epub 2011/11/11. doi: [10.1056/NEJMc1106975](https://doi.org/10.1056/NEJMc1106975) PMID: [22070496](https://pubmed.ncbi.nlm.nih.gov/22070496/).
19. Jensen TJ, Dzakula Z, Deciu C, van den Boom D, Ehrich M. Detection of microdeletion 22q11.2 in a fetus by next-generation sequencing of maternal plasma. *Clinical chemistry*. 2012; 58(7):1148–51. doi: [10.1373/clinchem.2011.180794](https://doi.org/10.1373/clinchem.2011.180794) PMID: [22563040](https://pubmed.ncbi.nlm.nih.gov/22563040/).
20. Srinivasan A, Bianchi DW, Huang H, Sehnert AJ, Rava RP. Noninvasive detection of fetal subchromosome abnormalities via deep sequencing of maternal plasma. *American journal of human genetics*. 2013; 92(2):167–76. doi: [10.1016/j.ajhg.2012.12.006](https://doi.org/10.1016/j.ajhg.2012.12.006) PMID: [23313373](https://pubmed.ncbi.nlm.nih.gov/23313373/); PubMed Central PMCID: [PMC3567270](https://pubmed.ncbi.nlm.nih.gov/PMC3567270/).
21. Yu SC, Jiang P, Choy KW, Chan KC, Won HS, Leung WC, et al. Noninvasive prenatal molecular karyotyping from maternal plasma. *PloS one*. 2013; 8(4):e60968. doi: [10.1371/journal.pone.0060968](https://doi.org/10.1371/journal.pone.0060968) PMID: [23613765](https://pubmed.ncbi.nlm.nih.gov/23613765/); PubMed Central PMCID: [PMC3629174](https://pubmed.ncbi.nlm.nih.gov/PMC3629174/).
22. Chu T, Yeniterzi S, Rajkovic A, Hogge WA, Dunkel M, Shaw P, et al. High resolution non-invasive detection of a fetal microdeletion using the GCREM algorithm. *Prenat Diagn*. 2014; 34(5):469–77. Epub 2014/01/24. doi: [10.1002/pd.4331](https://doi.org/10.1002/pd.4331) PMID: [24452987](https://pubmed.ncbi.nlm.nih.gov/24452987/).
23. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2013; 42(Database issue):D986–92. Epub 2013/11/01. doi: [gkt958](https://doi.org/gkt958) [pii] doi: [10.1093/nar/gkt958](https://doi.org/10.1093/nar/gkt958) PMID: [24174537](https://pubmed.ncbi.nlm.nih.gov/24174537/); PubMed Central PMCID: [PMC3965079](https://pubmed.ncbi.nlm.nih.gov/PMC3965079/).
24. Chu T, Bunce K, Hogge WA, Peters DG. A novel approach toward the challenge of accurately quantifying fetal DNA in maternal plasma. *Prenatal diagnosis*. 2010; 30(12–13):1226–9. Epub 2010/11/13. doi: [10.1002/pd.2656](https://doi.org/10.1002/pd.2656) PMID: [21072784](https://pubmed.ncbi.nlm.nih.gov/21072784/).