

RESEARCH

Open Access



Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data

Ruen Yao¹, Cheng Zhang², Tingting Yu¹, Niu Li¹, Xuyun Hu^{1,2}, Xiumin Wang³, Jian Wang¹ and Yiping Shen^{1,2*}

Abstract

Background: Whole exome sequencing (WES) has been widely accepted as a robust and cost-effective approach for clinical genetic testing of small sequence variants. Detection of copy number variants (CNV) within WES data have become possible through the development of various algorithms and software programs that utilize read-depth as the main information. The aim of this study was to evaluate three commonly used, WES read-depth based CNV detection programs using high-resolution chromosomal microarray analysis (CMA) as a standard.

Methods: Paired CMA and WES data were acquired for 45 samples. A total of 219 CNVs (size ranged from 2.3 kb – 35 mb) identified on three CMA platforms (Affymetrix, Agilent and Illumina) were used as standards. CNVs were called from WES data using XHMM, CoNIFER, and CNVnator with modified settings.

Results: All three software packages detected an elevated proportion of small variants (< 20 kb) compared to CMA. XHMM and CoNIFER had poor detection sensitivity (22.2 and 14.6%), which correlated with the number of capturing probes involved. CNVnator detected most variants and had better sensitivity (87.7%); however, suffered from an overwhelming detection of small CNVs below 20 kb, which required further confirmation. Size estimation of variants was exaggerated by CNVnator and understated by XHMM and CoNIFER.

Conclusion: Low concordances of CNV, detected by three different read-depth based programs, indicate the immature status of WES-based CNV detection. Low sensitivity and uncertain specificity of WES-based CNV detection in comparison with CMA based CNV detection suggests that CMA will continue to play an important role in detecting clinical grade CNV in the NGS era, which is largely based on WES.

Keywords: Clinical sequencing, Copy number variants, Whole exome sequencing, Structural variation

Background

Copy number variants are important human genomic variants known to be responsible for Mendelian disorders as well as for common genetic conditions such as autism, intellectual disability, and schizophrenia [1–3]. Chromosomal microarray analysis (CMA) has demonstrated its technical validity and has remained the method of choice for the detection of genome-wide copy number variants (CNVs) in clinical settings. It has also demonstrated its clinical validity for both pre- and

postnatal diagnostic testing [4, 5]. CMA is currently regarded as the gold standard for detection of CNVs that range from several kilobases to several megabases in size [6, 7].

The advent of next-generation sequencing (NGS) technology has dramatically improved our capability for examining small-scale sequence variants; it has also provided new options for the evaluation of large scale structural variants such as CNVs [8]. Whole-exome sequencing (WES) has been accepted as the most comprehensive test currently implemented in the clinical setting for small sequence variants [9, 10]. Much effort has been focused to generate CNV information from WES data [11]; however, low sensitivity and high false positive rates have been reported in previous studies using

* Correspondence: yiping.shen@childrens.harvard.edu

¹Department of Medical Genetics and Molecular Diagnostic Laboratory, Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai 200127, China

²Boston Children's Hospital, Boston, MA 02115, USA

Full list of author information is available at the end of the article

cancer cell lines [12], publicly available exome data [13], or comparing with whole genome sequencing data based CNV calling [14–16]. Thus, its technical validity has yet to be thoroughly evaluated.

Here, we evaluated three representative and popular read-depth based CNV detection programs: the eXome-Hidden Markov Model (XHMM), the Copy Number Inference From Exome Reads (CoNIFER), and CNVnator using clinical grade WES data. XHMM and CoNIFER detect rare CNVs based on a batched-comparison principle, while CNVnator detects CNVs based on a mean-shift approach within single samples. CNVs detected from the CMA platform were used as reference standard.

Methods

Samples and ethics statement

A total of 45 clinical diagnostic samples were enrolled from the Shanghai Children's Medical Centre and the Maternal and Child Health Hospital of the Guangxi Zhuang autonomous region with the approval of respective institutional ethics review committees. Genomic DNA was extracted using the QIAamp Blood DNA Mini kit® (Qiagen GMBH, Hilden, Germany).

WES and WES-based CNV detection

Exome targets were captured using the Agilent SureSelect Human All Exon V4 or V5 kit (Agilent Technologies, Santa Clara, CA). Raw sequencing data (FASTQ format) were generated via the Illumina HiSeq 2000 platform (Illumina, Inc., San Diego, CA). The Burrows Wheeler Alignment tool (BWA) v0.2.10 [17] was employed for sequencing data alignment to the Human Reference Genome (NCBI build 37, hg 19). All data were assessed using FastQC (version 0.11.2) (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) for quality.

CNVs were generated using the following three CNV detection programs: (1) XHMM v1.0 [18], (2) CoNIFER v0.2.2 [19], and (3) CNVnator v0.2.7 [20]. XHMM includes several analytic steps and involves a number of parameters. In our study, we set all parameters to default (minTargetSize: 10; maxTargetSize: 10,000; minMeanTargetRD: 10; maxMeanTargetRD: 500; minMeanSampleRD: 25; maxMeanSampleRD: 20; maxSdSampleRD: 150) for filtering samples and targets, and prepared the data for normalization via XHMM. The only parameter that could be adjusted on Conifer was SVD, which was set to 1. For CNVnator, we set the bin size to 50–60 according to the average coverage depth of our sequencing data (45–70 X). XHMM and CoNIFER used a pooled sample calling approach as input, and CNVnator called CNVs sample by sample after individually generating a baseline.

CMA and CMA-based CNV detection

CMA were performed using three different array platforms including the SurePrint G3 customized array (Agilent Technologies, Santa Clara, CA), CytoScan HD (Affymetrix, Santa Clara, CA), and Infinium iSelect HD and HTS Custom Genotyping BeadChips (Illumina, San Diego, CA). Prior validated settings for each platform were consistently utilized for CNV detection and filtering. CNVs in the size range of 2 kb – 400 kb were detected via CMA and were further confirmed by manual inspection.

Results

Quality control of WES data

Fourteen samples were prepared using the Agilent SureSelect Human All Exon V4 kit and the remaining samples were prepared using the V5 kit. The mean read depth of all samples ranged around 50 X and the average read quality was well above the standard of 20 X. Details of sequence data are available in the supplemental data (Additional file 1: Table S1).

Size distribution of CNV detected via CMA and WES

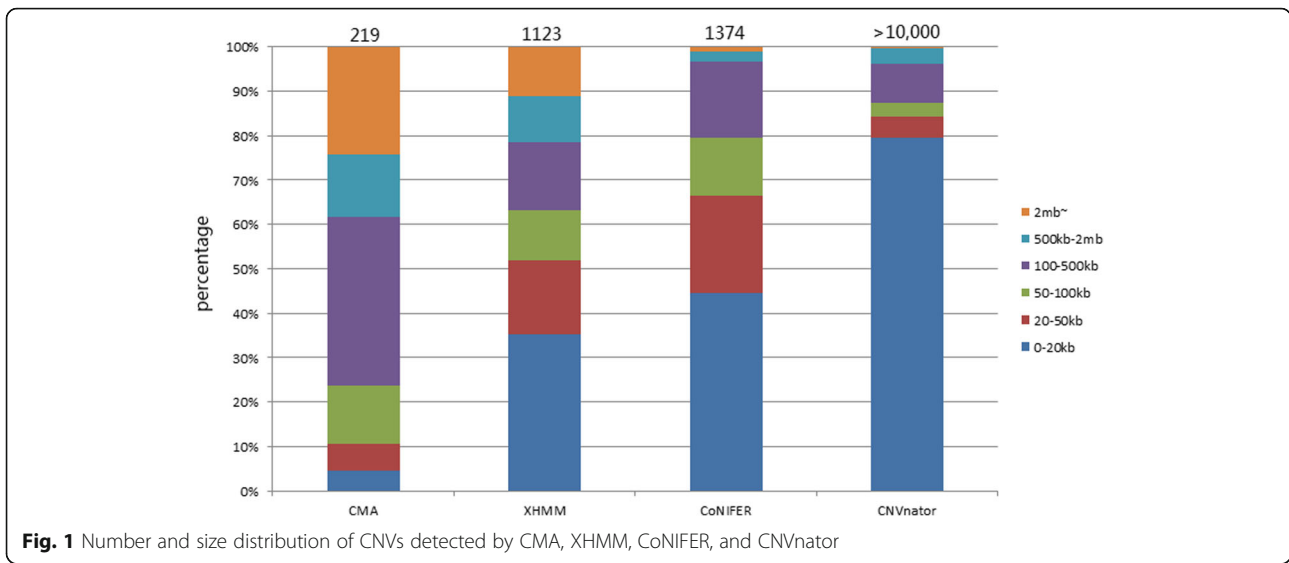
A total of 219 CNVs were detected via CMA from all samples. Forty-eight CNVs were located in regions that had no exome capture probes; consequently, they were removed from being used as true CNVs when comparing data between CMA and NGS. The remaining 171 CNVs were in regions involving at least one exon. The CNVs were examined and compared for size distribution, detection sensitivity, boundaries, and overlap among three programs and between two platforms.

1. CNV size

We arbitrarily constructed six size bins as shown in Fig. 1. The largest portion (37.9%) of CNV detected by CMA ranged within 100–500 kb whereas CNV detected by NGS data were of much smaller size; 35.3, 44.5 and 79.5% of CNVs were detected by XHMM, CoNIFER, and CNVnator, respectively and belong to the 0–20 kb bin. CNVnator in particular detected many smaller CNVs (42.2% below 10 kb, 27.3% below 5 kb).

2. Detection sensitivity

We defined the detection of any particular CNV when there was a 50% overlap with a CNV detected via CMA. Using this definition for the presence/absence of CNV, 25, 38, and 150 CNVs were found to be detected by CoNIFER, XHMM, and CNVnator respectively; thus, the detection sensitivities of three programs were 14.6, 22.2 and 87.7%, respectively. CoNIFER and XHMM have an even poorer detection sensitivity for smaller CNVs involving fewer capturing probes, whereas CNVnator had



a rather consistent detection sensitivity for all size CNVs (> 3 kb) (Fig. 2).

3. Precision of CNV detection

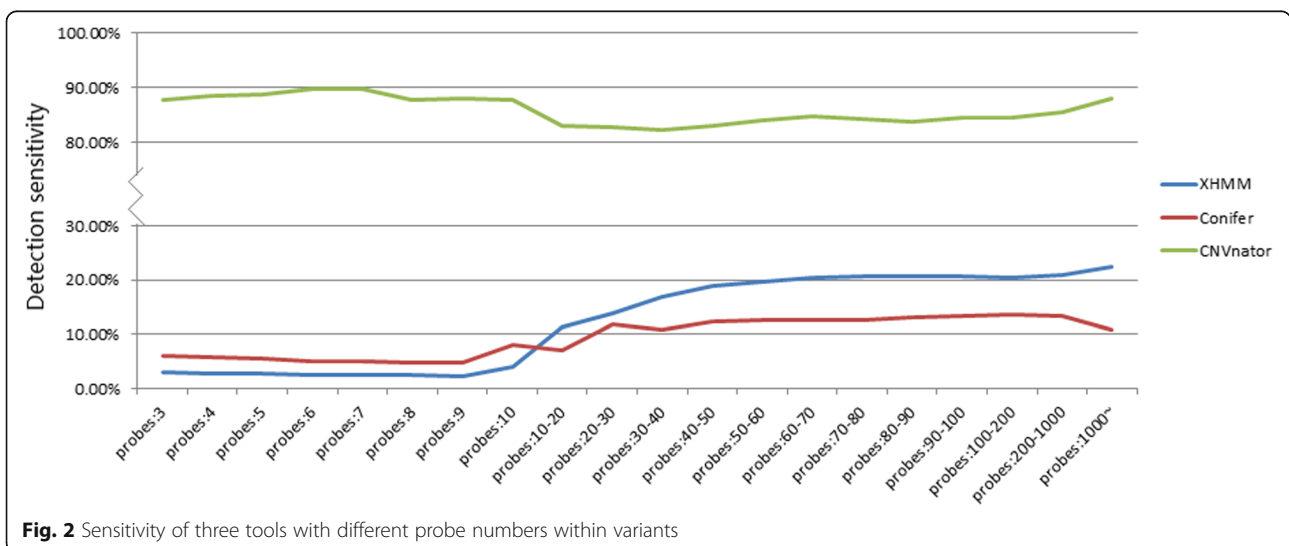
Among the 171 variants detected by CMA, 152 variants were detected by at least one WES program. Forty-six variants were detected by XHMM and CoNIFER, which are shown in Fig. 3(A). We plotted the size ratio of those detected from three programs, using CMA as reference. XHMM and CoNIFER detected more accurate size of variants, while CNVnator reported a significantly larger CNV size (Fig. 3(B)).

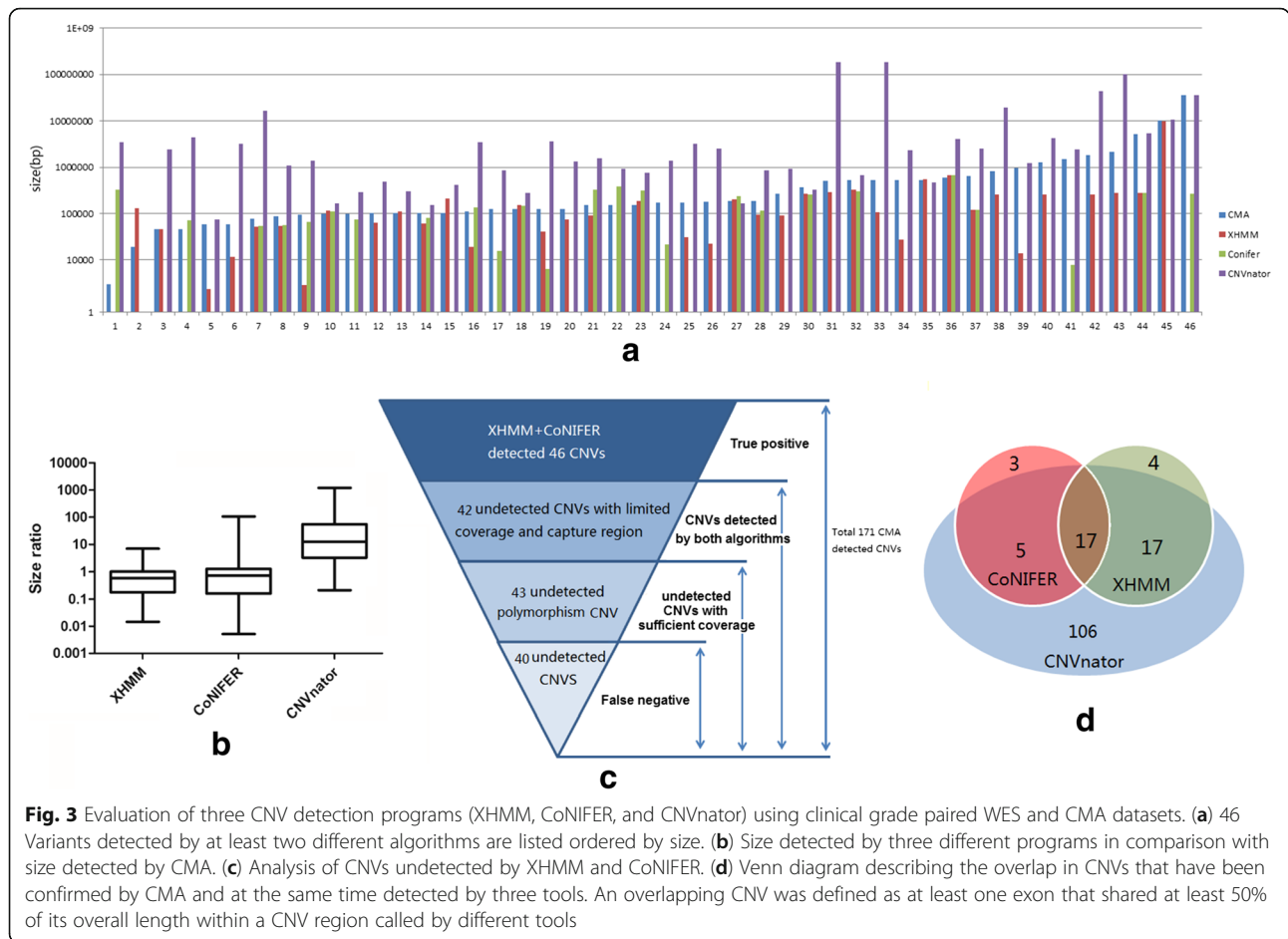
4. Characteristics of CNV missed by exome data (Fig. 3 (C))

A large number of CNVs (125) were missed by CoNIFER and XHMM combined detection and were further investigated. The WES read coverage and capture probes distribution were insufficient for both CoNIFER and XHMM detection of 42 CNVs. XHMM and CoNIFER automatically filtered capturing probes located in the region within recurrent CNV detected in the same batch; thus, 43 variants were missed and had to be confirmed with involving probes. Details of all 171 variants are available in the supplemental data (Additional file 2: Table S2).

5. Poor concordance among three programs (Fig. 3 (D))

Although CoNIFER and XHMM used a similar batched input approach, poor CNV detection concordance was





still identified in our study. CNVnator discovered most variants and covered most variants that could be detected through WES data. Only 17 CNVs were detected by all three programs.

6. Detection of clinical relevant variants

All variants were evaluated with our in-house standard for clinical relevant variants and eight CNVs were categorized as pathogenic or likely pathogenic variants ranging from 306 kb – 35 mb. Six of these variants were detected by WES programs. A 306 kb variant on chromosome 3 remained undetected due to particularly low capture probe coverage within the variant region. Another 11 mb variant on chromosome 2 remained undetected despite sufficient capture probes and depth coverage (Additional file 3: Table S3).

Discussion

Copy number variants (CNVs) are a very important target in the clinical diagnosis of genetic diseases. CMA has been proven as the most stable and accurate platform for CNV detection and has been implemented as a

clinical test for more than a decade. NGS now provided a new approach for detecting CNV, which can potentially replace CMA. Before implementing NGS-based CNV detection, extensive validation is required to evaluate the validity of the new method.

Numerous WES based CNV detection programs have been developed, including the 15 read-depth based CNV detection tools currently available [21]. We selected three representative and well-known methods for this study. XHMM is the most commonly accepted software, which employs the classical hidden Markov model (HMM) for CNV identification and achieves a sensitivity of 8–14% via XHMM, reported against CNV detection based on WGS data [13]. The XHMM framework starts with aligned BAM files to calculate the depth of coverage; then, utilizing normalized read depths via principal component analysis (PCA). Finally, XHMM uses the normalized data to train and run a Hidden Markov Model (HMM) for CNV detection. CoNIFER was the first developed tool to deal with rare CNVs from multiple samples and has been chosen as representative software, which can be used as reference in evaluating other new softwares [22]. CoNIFER calculates the RPKM

(reads per kilobase per million mapped reads) values for each sample, and utilizes the singular value decomposition (SVD) method (originating from linear algebra) to reduce data dimensions for detecting obvious CNV signals. Evaluation against the array CGH platform in breast cancer samples characterized CoNIFER as leading to high false positives, low sensitivity, and obvious duplication bias [11]. Another study showed that CoNIFER achieves higher precision, but at a cost of reduced sensitivity below 5% [13]. XHMM and CoNIFER have been evaluated in parallel in patients with nonsyndromic hearing loss showing poor concordance on size of detected CNV [23]. However, both tools are noted for advantages of identification of rare CNV from a population of WES samples [24]. CNVnator was previously used in whole genome data for CNVs identification based on read depth, and was accessed to achieve better resolution of CNV borders than the other WGS data-based tools [25]. The main methodology for CNVnator is a mean-shift. The software first divides the whole genome into equal sized, non-overlapping bins, and treats the mapped reads of each bin as a read depth signal. To estimate copy number change in each genome segment, it then calculates the *P*-value for a one-sample t-test, testing whether the mean RD signal of a segment would be close to the genome average. In a comprehensive comparison study, CNVnator was accessed to be outstanding in break point position and copy number estimation; however, discordance of variants was also discovered among all tools evaluated in the study [26].

In our study, large differences were observed in number and size distribution of CNVs detected from CMA and three WES based tools. Microarray platforms have a smaller capacity to detect small variants that are not covered by a sufficient number of probes. Several studies have tried to understand the roles of these small variants. The detection of small, non-recurrent pathogenic or likely pathogenic CNVs could help to increase the diagnostic yield of CMA clinical testing by ~3% [27, 28]. WES-based tools, such as XHMM and CoNIFER, are capable of detecting small variants as long as a sufficient number of capturing probes (> 10) are covered in the region and enable a sensitivity of 14.6 and 22.2%, respectively, indicating the importance of probe number for CNVs detection. The overwhelming number of variants CNVnator detected from samples was due to the extreme resolution of the algorithm [19]. This extreme resolution is affected by sequencing depth and high resolution could result in splitting large CNVs into small pieces, which are more sensitive in detecting smaller variants. Larger bin size setting in CNVnator could help to merge consecutive small CNVs as integrated variants; however, this parameter was limited by the average sequencing depth of our clinical WES data.

125 CMA confirmed CNVs that were not detected by XHMM and CoNIFER were further investigated for possible explanations. Low sequencing depth (< 10 X) and limited capture probes (< 10) were detected in 42 variants and these regions were automatically excluded during the normalization step of both tools. The detection for these CNV may be improved if sequence depth increased. The programs filter out capture probes located in recurrent variants that detected the same batch during data processing; thus, 43 polymorphism CNVs were neglected during the detection, which was also confirmed by our in-house array database [http://database.gdg-fudan.org/DB_HTML/DataSub.html]. Thus, only 40 (23.4%) CNVs remained theoretically undetected. Limitation of sample number and sequencing depth of XHMM and CoNIFER could be a possible explanation of these undetected variants. CoNIFER requires at least 50 million mapped reads and a minimum of eight exome samples to run at a time, while XHMM recommends ~50 exome samples with at least 60–100 X coverage [18, 29]. Characteristics of samples in each batch also contribute to the effectiveness of CNV detection in XHMM and CoNIFER. Recurrent pathogenic or likely pathogenic variants may be filtered out erroneously, if they existed in multiple samples; therefore, including non-abnormal reference samples as part of the batch could help to detect these CNVs. Conservative predefined thresholds in default settings of the CoNIFER might be a further reason for missing variants. Read-depth based tools are fairly limited to repeated regions of the reference genome [30]; thus, the sequence nature of specific locations also hinders detection of variants. CNVnator was designed for CNV discovery and genotyping from read-depth analysis based on a mean-shift approach. The number of nucleotides covered in each shift is called bin size (50–60 in our study), which can be determined by the average coverage of sequencing data (45–70 for our samples). CNVnator had the highest sensitivity of 87.7% since 150 of 171 CMA confirmed variants were detected by CNVnator. A Venn diagram was used to show the poor discordance among WES based tools, which was attributed to unsatisfying sequencing depth and inadequate number of batched samples. Therefore, CNV detection from WES based tools was affected by the following factors (ranked in the order of importance): probe number, reading depth, sample constituent in the batch, software parameters, and sequence nature of variants.

Using CMA detected variants as standard, the three tested WES based CNV detecting tools were not able to detect the accurate size of variants from WES data. XHMM and CoNIFER have lower sensitivity, but more accurate size of CNVs compared to CNVnator. CNVnator reached higher sensitivity at the cost of high false

positive rates and exaggerated readout of the variant size. Poor concordance of CNV detection was observed in the study. Increasing the number of batch samples and valid sequencing depth were the most realizable approaches to improve performance of these WES based tools. At this stage, CMA still remains the first-choice and gold standard for CNV detection for clinical diagnostic purpose. CNV detection tools using WES data could be used as a screening tool.

Conclusion

Low concordances of CNV detection were observed via three different read-depth based programs indicating that WES-based CNV detection still remains immature and unstable compared to CMA. Since WES based CNV detection was evaluated to have low sensitivity and uncertain specificity in comparison with CMA based CNV detection, CMA will continue to play an important role in detecting clinical grade CNV in the NGS era, which is largely based on WES. CNV detection tools using WES data could be considered as a complementary way with only computational effort, but where further validation has been suggested for the purpose of clinical diagnosis.

Additional files

Additional file 1: Details and quality control information of sequencing data. (XLS 32 kb)

Additional file 2: Details of all 171 variants detected by CMA. (XLS 90 kb)

Additional file 3: Details of eight clinical relevant CNVs. (XLS 37 kb)

Acknowledgements

The authors would like to thank all members of family for their participation in this study.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 81371903 and 81,472,051), the Project of Shanghai Municipal Science and Technology Commission (Grant No. 15410722800), and the Project of Shanghai Municipal Education Commission- Gaofeng Clinical Medicine (Grant No. 20152529).

Availability of data and materials

All data that were generated or analyzed during this study are included in this publication and its supplemental data. Please contact the corresponding author for further data requests.

Authors' contributions

RY was the major contributor in writing the manuscript and analyzing the data; CZ, TY, and XH contributed to the computational analysis of CMA data and WES data; NL and XW helped in summarizing data and interpretation; JW and YS performed a critical review of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study was approved by the Committee on Ethics of the Shanghai Children's Medical Centre and the Maternal and Child Health Hospital of the Guangxi Zhuang autonomous region. Furthermore, we obtained written informed consent from the patients, performing all the experiments according to the regulations in the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Medical Genetics and Molecular Diagnostic Laboratory, Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai 200127, China. ²Boston Children's Hospital, Boston, MA 02115, USA. ³Department of Endocrinology and Metabolism, Shanghai Children's Medical Center, Shanghai Jiaotong University School of Medicine, Shanghai 200127, China.

Received: 26 April 2017 Accepted: 15 August 2017

Published online: 23 August 2017

References

1. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet.* 2006;7(2):85–97.
2. Sharp AJ, Cheng Z, Eichler EE. Structural variation of the human genome. *Annu Rev Genomics Hum Genet.* 2006;7:407–42.
3. Martin CL, Kirkpatrick BE, Ledbetter DH. Copy number variants, aneuploidies, and human disease. *Clin Perinatol.* 2015;42(2):227–42. vii
4. Fiorentino F, Napoletano S, Caiazzo F, Sessa M, Bono S, Spizzichino L, Gordon A, Nuccitelli A, Rizzo G, Baldi M. Chromosomal microarray analysis as a first-line test in pregnancies with a priori low risk for the detection of submicroscopic chromosomal abnormalities. *Eur J Hum Genet.* 2013;21(7):725–30.
5. Manning M, Hudgins L; professional practice and guidelines committee. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. *Genet Med.* 2010;12(11):742–5.
6. Liang D, Peng Y, Lv W, Deng L, Zhang Y, Li H, Yang P, Zhang J, Song Z, Xu G, Cram DS, Wu L. Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J Mol Diagn.* 2014;16(5):519–26.
7. Boone PM, Bacino CA, Shaw CA, Eng PA, Hixson PM, Pursley AN, Kang SH, Yang Y, Wiszniewska J, Nowakowska BA, del Gaudio D, Xia Z, Simpson-Patel G, Immken LL, Gibson JB, Tsai AC, Bowers JA, Reimschisel TE, Schaaf CP, Potocki L, Scaglia F, Gambin T, Sykulski M, Bartnik M, Derwinska K, Wisniewicka-Kowalnik B, Lalani SR, Probst FJ, Bi W, Beaudet AL, Patel A, Lupski JR, Cheung SW, Stankiewicz P. Detection of clinically relevant exonic copy-number changes by array CGH. *Hum Mutat.* 2010;31(12):1326–42.
8. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, Chinwalla A, Conrad DF, Fu Y, Grubert F, Hajirasouliha I, Hormozdiari F, Iakoucheva LM, Iqbal Z, Kang S, Kidd JM, Konkel MK, Korn J, Khurana E, Kural D, Lam HY, Leng J, Li R, Li Y, Lin CY, Luo R, Mu XJ, Nemes J, Peckham HE, Rausch T, Scally A, Shi X, Stromberg MP, Stütz AM, Urban AE, Walker JA, Wu J, Zhang Y, Zhang ZD, Batzer MA, Ding L, Marth GT, McVean G, Sebat J, Snyder M, Wang J, Ye K, Eichler EE, Gerstein MB, Hurler ME, Lee C, SA MC, Korb JO, 1000 Genomes Project. Mapping copy number variation by population-scale genome sequencing. *Nature.* 2011;470(7332):59–65.
9. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461(7261):272–6.
10. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet.* 2014;59(1):5–15.
11. Miyatake S, Koshimizu E, Fujita A, Fukai R, Imagawa E, Ohba C, Kuki I, Nukui M, Araki A, Makita Y, Ogata T, Nakashima M, Tsurusaki Y, Miyake N, Saito H, Matsumoto N. Detecting copy-number variations in whole-exome sequencing data using the eXome hidden Markov model: an 'exome-first' approach. *J Hum Genet.* 2015;60(4):175–82.
12. Guo Y, Sheng Q, Samuels DC, Lehmann B, Bauer JA, Pietenpol J, Shyr Y. Comparative study of exome copy number variation estimation tools using

- array comparative genomic hybridization as control. *Biomed Res Int.* 2013;2013:915636.
13. Samarakoon PS, Sorte HS, Kristiansen BE, Skodje T, Sheng Y, Tjønnfjord GE, Stadheim B, Stray-Pedersen A, Rødningen OK, Lyle R. Identification of copy number variants from exome sequence data. *BMC Genomics.* 2014;15:661.
 14. Tan R, Wang Y, Kleinstejn SE, Liu Y, Zhu X, Guo H, Jiang Q, Allen AS, Zhu M. An evaluation of copy number variation detection tools from whole-exome sequencing data. *Hum Mutat.* 2014;35(7):899–907.
 15. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, Shang L, Boisson B, Casanova JL, Abel L. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A.* 2015;112(17):5473–8.
 16. Hehir-Kwa JY, Pfundt R, Veltman JA. Exome sequencing and whole genome sequencing for the detection of copy number variation. *Expert Rev Mol Diagn.* 2015;15(8):1023–32.
 17. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
 18. Fromer M, Moran JL, Chambert K, Banks E, Bergen SE, Ruderfer DM, Handsaker RE, McCarroll SA, O'Donovan MC, Owen MJ, Kirov G, Sullivan PF, Hultman CM, Sklar P, Purcell SM. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet.* 2012;91(4):597–607.
 19. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP; NHLBI Exome Sequencing Project, Quinlan AR, Nickerson DA, Eichler EE. Copy number variation detection and genotyping from exome sequence data. *Genome Res* 2012;22(8):1525–1532.
 20. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* 2011;21:974–84.
 21. Kadalayil L, Rafiq S, Rose-Zerilli MJ, Pengelly RJ, Parker H, Oscier D, Strefford JC, Tapper WJ, Gibson J, Ennis S, Collins A. Exome sequence read depth methods for identifying copy number changes. *Brief Bioinform.* 2015;16(3):380–92.
 22. Bansal V, Dorn C, Grunert M, Klaassen S, Hetzer R, Berger F, Sperling SR. Outlier-based identification of copy number variations using targeted resequencing in a small cohort of patients with Tetralogy of Fallot. *PLoS One.* 2014;9(1):e85375.
 23. Bademci G, Diaz-Horta O, Guo S, Duman D, Van Booven D, Foster J 2nd, Cengiz FB, Blanton S, Tekin M. Identification of copy number variants through whole-exome sequencing in autosomal recessive nonsyndromic hearing loss. *Genet Test Mol Biomarkers.* 2014;18(9):658–61.
 24. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics.* 2013;14(Suppl 11):S1.
 25. Legault MA, Girard S, Lemieux Perreault LP, Rouleau GA, Dubé MP. Comparison of sequencing based CNV discovery methods using monozygotic twin quartets. *PLoS One.* 2015;10(3):e0122287.
 26. Duan J, Zhang JG, Deng HW, Wang YP. Comparative studies of copy number variation detection methods for next-generation sequencing technologies. *PLoS One.* 2013;8(3):e59128.
 27. Hollenbeck D, Williams CL, Drazba K, Descartes M, Korf BR, Rutledge SL, Lose EJ, Robin NH, Carroll AJ, Mikhail FM. Clinical relevance of small copy-number variants in chromosomal microarray clinical testing. *Genet Med.* 2017;19(4):377–85.
 28. Poultney CS, Goldberg AP, Drapeau E, Kou Y, Harony-Nicolas H, Kajiwara Y, De Rubeis S, Durand S, Stevens C, Rehnström K, Palotie A, Daly MJ, Ma'ayan A, Fromer M, Buxbaum JD. Identification of small exonic CNV from whole-exome sequence data and application to autism spectrum disorder. *Am J Hum Genet.* 2013;93(4):607–19.
 29. Fromer M, Purcell SM. Using XHMM Software to detect copy number variation in whole-exome sequencing data. *Curr Protoc Hum Genet.* 2014;81:7.23.1–21.
 30. Alkan C, Kidd JM, Marques-Bonet T, Aksay G, Antonacci F, Hormozdiari F, Kitzman JO, Baker C, Malig M, Mutlu O, Sahinalp SC, Gibbs RA, Eichler EE. Personalized copy number and segmental duplication maps using next-generation sequencing. *Nat Genet.* 2009;41(10):1061–7.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

