AMERICAN JOURNAL OF MEDICAL AND CLINICAL SCIENCES, 2023
VOL 8, NO. 4, PAGE 1 – 2

RESEARCH ARTICLE

Open Access

Validating Clinical Exome Sequencing (CES) Test for Copy Number Variation (CNV) Analysis in Patients with Autism Spectrum Disorder

Katia Margiotti[1]*, Marco Fabiani[1], Alvaro Mesoraca[1] and Claudio Giorlandino[1,2]

[1]Human Genetics Lab, Altamedica Main Centre, Viale Liegi 45, 00198 Rome, Italy.
[2]Department of Prenatal Diagnosis, Altamedica, Fetal-Maternal Medical Centre, Viale Liegi 45, 00198 Rome, Italy.

ABSTRACT

A routine methodologies to detect germline copy number variants (CNVs) are ArrayCGH, and SNParray, while clinical exome sequencing (CES) it is not a standard methods used in clinical settings. This study aims to investigate the efficiency of CNVs identification by CES in the diagnosis of Autism spectrum disorder (ASD) based on a Italian cohort. In this cohort study, 50 patients with ASD diagnosis were recruited for ASD screening. Simultaneous CNVs and single nucleotide variation (SNV) analysis was conducted to identify the anomalies that can cause ASD phenotype. A total of 13 patients were identified with a causative CNVs among 15 (86.67%) SNP array positive patients. Showing that CES data can be used for detecting clinically relevant CNVs with high sensitivity for use in clinical diagnostic settings.

ARTICLE HISTORY

Received 10 Oct 2023
Accepted 19 Nov 2023
Published 26 Nov 2023

KEYWORDS

Clinical exome sequencing (CES), Copy number variants (CNVs), Autism spectrum disorder (ASD).

Introduction

Autism spectrum disorder (ASD) is a complex and genetically heterogeneous disorder. In the last decades, the remarkable advances in the knowledge of genetic aspects of autism based on genetic and molecular research, as well as the development of new molecular diagnostic tools, have substantially improve the genetic diagnoses. Nowadays, it is estimated that using the currently available molecular tests, a potential underlying genetic cause can be identified in nearly 40% of cases [1-3]. Therefore, in view of the current knowledge about the genetic architecture of autism spectrum disorder, molecular genetic investigation has become increasingly important. Currently, a number of specific genetic variant are known to be associated with ASD. The chromosomal copy number variants (CNV), such as microdeletions and microduplication could be associated with ASD in about 15-20% of patients [4]. Clinical exome sequencing (CES) has been widely utilized to investigate disruptive single nucleotide variants (SNV) associated ASD in about 8-20% of individuals [5,6]. CNVs detection using CES data is not currently a routine clinical test, likely due to the overwhelming inconsistencies among different methods [7-9] and the lack of a high-quality reference for CNVs from CES data. In this study, we used a cohort of 50 samples with clinical SNP array data and CES data to estimate the sensitivity and false discovery rate of the CNVs detection using clinical exome sequencing. Our results show that CES data can be used for detecting clinically important CNVs with high sensitivity and for use in clinical diagnostic settings.

Results and Discussion

A total of 50 SNP array analysis of patients with autism spectrum disorder (ASD) were evaluated for validate ours Clinical exome sequencing (CES) test results. The clinical presentations (phenotypes) were categorized to ASD-Simplex (no dysmorphic features), ASD-Complex (with dysmorphic features) by a psychologist (Table 1). The median age in this testing cohort was 6 years old. The male percentage was 68% versus the 32% of female (Table 1). In this cohort, the overall frequency CNV detection was 3 % (15/50) (Table 2). To assess the accuracy of CES based CNV detection pipeline results were selected 15 SNP array positive patients (9 ASD-Complex, 6 ASD-Simplex) and 35 SNP array negative patients (28 ASD-Simplex, 7 ASD-Complex). Based on the confirmatory SNP array, the positive predictive value (PPV) indicating the probability that samples with a positive test truly have the genetic variants, was available for 15 samples with microdeletion or duplication in several regions of the genome. Among the 15 cases with CNVs were found to be true positive 13 cases, and to be false negative 2 cases. The estimated PPV for CNVs was 81,25% (Table 2). Based on the negative SNP array, the negative predictive value (NPV) indicating the probability that samples with a negative test would actually not have the genetic variants was available for 35 samples that resulted negative after SNP arrays analysis (Table 2). Three of the 35 negative samples return a positive result, giving us an NPV of 94.12%. CES analysis achieved 86.67% sensitivity and 91.43% specificity, in comparison to the gold standard for CNVs detection the SNP array analysis (Table 2). Using a cohort of 50 samples, we validated an CES based CNV detection pipeline using a set of true positive (n=15) and negative (n=35) CNVs from SNP array. Compared to the SNP array, the CES based CNV detection pipeline had an 87% true-positive rate. The pipeline had a false discovery rate of 8,5% compared to the SNP array. Detecting CNVs from ES data is perceived as challenging for clinical use as many previously published reports suggested high false-positive rates and low

Contact Katia Margiotti, Human Genetics Lab, Altamedica Main Centre, Viale Liegi 45, 00198 Rome, Italy, Tel: +39 06 8505805, Fax: +39 068505815.

sensitivity. For use in a clinical setting, one requires the highest sensitivity and lowest false discovery rate possible for a given platform. In this work, we used a cohort of 50 individuals to validate ours Clinical exome sequencing (CES) test in the ability to detect CNVs from exome sequencing data obtained 86.67% sensitivity and 8,5% false discovery rate. Secondary evaluation of the raw date of the 2 false negative samples reveals low quality of the sequencing output. The 3 false positives were all monoexonic genes deletion, suggesting that CNV calls encompassing genes without introns may also require further assessments. In conclusion, it is essential to maintain investments in resources and continue to implement the tools that are already available in order to make NGS data a better diagnostic test overall.

Methods

The study includes the investigation of 50 patients who was admitted at Altamedica Medical Centre (Rome, Italy) and undergone to SNParray and Clinical exome sequencing (CES) diagnostics test. The study was approved by local ethical committee and all participating were provided with written informed consent. Genomic DNA was extracted from peripheral blood using the DNeasy Blood & TissueKit and QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Cases were genotyped using the single nucleotide polymorphism (SNP)-based array "CytoSNP-850K v1.3 BeadChip". This array contains 850.000 SNPs that are evenly distributed across the genome. All chromosomal positions are given in human genome build hg19 (GRCh37 Genome Reference Consortium Human Build 37 (GRCh37)). Clinical exome sequencing (CES) was carried out using the TruSight One Sequencing Panel (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. The panel covers 4813 disease-associated genes. The targeted exonic regions underwent paired-end sequencing on an Illumina platform, using a NextSeq 550Dx sequencing system (Illumina, San Diego, CA, USA). The detected variants were annotated and filtered, based on the information of the functional prediction tools [e.g., Polyphen2, SIFT, REVEL) and public disease variant database (e.g., ClinVar, HGMD, OMIM, dbSNPs and GWAS). CNV calls were performed using the DRAGEN Enrichment CNV calling module with 50 healthy sample as baseline. Each Variant Call Format (VCF) file was then annotated adding crucial metadata for interpretation by the clinician using enGenome software. The detected Structural Variants (SV) were always visualized and manually checked in Integrative Genomics Viewer (IGV).

Table 1: Patient characteristics undergoing to SNP array clinical testing for ASD.

Clincal charateristics	Value
Median Age	6 (3-50)
Male	29 (58%)
Famale	21 (42%)
Indications for SNParray	**Ratios %**
ASD-Simplex	34 (68%)
ASD-Complex	16 (32%)

Table 2. Efficiency of Clinical exome sequencing (CES) test for CNVs detection.

	SNP array (Positive)	SNP array (Negative)	
Total =50			
NGS (Positive)	**13**	**3**	
NGS (Negative)	**2**	**32**	
Specificity %	**Sensitivity %**	**PPV%**	**NPV%**
91.4 (95% CI, 76.94 to 98.20)	**86.67** (95% CI, 59.54-98.34)	**81.25** (95% CI 54.35 to 95.95)	**94.12** (95% CI 80.32 to 99.28)

PPV=positive predictive value, NPV=negative predictive value

Authors' contributions

All authors have materially participated in the study and manuscript preparation. KM, MF, carried out all the molecular genetic analysis, and participated in the design of the work; KM and MF collected all clinical data and participated in conceiving the manuscript; AM and CG participated in conceiving the study, and drafted and revised the manuscript. All authors have approved the final article.

Conflict of interest

KM, MF, and AM, are employed by Altamedica Medical centre, and CG is the scientific director of Altamedica Medical centre of Rome.

References

1. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron. 2015; 87: 1215-1233.
2. Tammimies K, Marshall CR, Walker S, Kaur G, Thiruvahindrapuram B, et al. Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. JAMA. 2015; 314: 895-903.
3. Schaefer GB, Mendelsohn NJ. Professional Practice and Guidelines Committee Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. Genet Med. 2013; 15: 399-407.
4. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, et al. Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet. 2008; 82: 477-488.
5. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. J Mol Diagn. 2018; 20: 4-27.
6. Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, et al. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. Nat Biotechnol. 2014; 32: 246-251.
7. Yao R, Zhang C, Yu T, Li N, Hu X, et al. Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. Mol Cytogenet. 2017; 10: 30.
8. Kim HY, Choi JW, Lee JY, Kong G. Gene-based comparative analysis of tools for estimating copy number alterations using whole-exome sequencing data. Oncotarget. 2017; 8: 27277-27285.
9. Hong CS, Singh LN, Mullikin JC, Biesecker LG. Assessing the reproducibility of exome copy number variations predictions. Genome Med. 2016; 8: 82.