BMC Medical Genomics

RESEARCH

Open Access

Diagnostic yield of exome sequencing-based copy number variation analysis in Mendelian disorders: a clinical application

Tahir Atik[1], Enise Avci Durmusalioglu[1*], Esra Isik[1], Melis Kose[2], Seda Kanmaz[3], Ayca Aykut[4], Asude Durmaz[4], Ferda Ozkinay[1] and Ozgur Cogulu[1]

Abstract

Next-generation sequencing (NGS) coupled with bioinformatic tools has revolutionized the detection of copy number variations (CNVs), which are implicated in the emergence of Mendelian disorders. In this study, we evaluated the diagnostic yield of exome sequencing-based CNV analysis in 449 patients with suspected Mendelian disorders. We aimed to assess the diagnostic yield of this recently utilized method and expand the clinical spectrum of intragenic CNVs. The cohort underwent whole exome sequencing (WES) and clinical exome sequencing (CES). Using GATK-gCNV, we identified 12 pathogenic CNVs that correlated with their clinical findings and resulting in a diagnostic yield of 2.67%. Importantly, the study emphasizes the role of CNVs in the etiology of Mendelian disorders and highlights the value of exome sequencing-based CNV analysis in routine diagnostic processes.

Keywords Copy number variations, Next-generation sequencing, CNV analysis, Mendelian disorders

Introduction

Copy number variations (CNVs), including deletions and duplications within the genome, play a pivotal role in genetic diversity and disease development. Detecting CNVs is crucial for diagnosing a variety of disorders [1]. Traditional techniques such as microarrays and Multiplex Ligation-dependent Probe Amplification (MLPA) have commonly been used for CNV detection, but the introduction of next-generation sequencing (NGS) has revolutionized the field due to its efficiency and cost-effectiveness [2–4].

The microarray has been a powerful tool for CNV detection in the human genome. However, it has limitations, particularly in resolution. While it can accurately detect CNVs of a certain size depending on the method used, its ability to identify smaller variants, such as exonic deletions and duplications, is limited [5]. Conversely, MLPA is a valuable technique for targeting specific genes and regions of interest. As a PCR-based technology, it uses probes to selectively bind and amplify targeted sequences. However, it requires the design and synthesis of specific probes for each gene or region of interest [6].

CNV detection can be accomplished by analyzing the depth of coverage of short reads generated by NGS. This process involves dividing the genome into specific regions, or 'bins' (e.g., 100 base pairs), predetermined by the laboratory. After sequencing, read alignment,

*Correspondence:
Enise Avci Durmusalioglu
eniseavci.ea@gmail.com
[1]Pediatric Genetics, Ege University, Izmir, Turkey
[2]Department of Genetics, The Children Hospital of Philadelphia, Mitochondrial Medicine Frontier Program, Philadelphia, USA
[3]Pediatric Neurology, Ege University, Izmir, Turkey
[4]Medical Genetics, Ege University, Izmir, Turkey

data filtering, and normalization, the number of reads in each bin of the patient sample is compared with the corresponding bin in a reference set composed of normal samples. Sophisticated algorithms are employed to detect deletions and duplications by comparing the read depths. The resulting data, which is used to calculate relative copy number information, can be visualized and interpreted through a user interface for further analysis [4, 7].

In this study, we evaluated 12 patients diagnosed using NGS-based CNV analysis. Our aim was to underline the role of intragenic deletions and duplications in the etiology of Mendelian disorders and evaluate the diagnostic yield of NGS-based CNV analysis. By using this advanced genomic approach, we sought to gain insights into the genetic mechanisms underlying these disorders and enhance our understanding of the clinical implications of CNVs.

Materials and methods

Ethical approval

The present study was approved by the Ethics Committee of Ege University Faculty of Medicine (Reference No: 23-7T/27). All procedures involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with Helsinki Declaration. Written informed consent was obtained from the parents or guardians of the participants, ensuring transparency and adherence to privacy regulations.

Patients

A total of 215 whole exome sequencing (WES) analyses and 234 clinical exome sequencing (CES) analyses were performed at the Ege University Faculty of Medicine, Department of Pediatric Genetics, from September 1, 2021, to November 1, 2022. These analyses included a cohort of patients aged 1–18 years, all of whom presented with diverse clinical findings suggestive of monogenic disorders. The exome sequencing-based CNV analysis identified 12 patients with a molecular etiology. The cohort included patients with various genetic indications, such as cardiomyopathy, lactic acidosis, spastic paraplegia, neuronal ceroid lipofuscinosis, epidermolysis bullosa, unidentified vision loss with deafness, and preliminary diagnoses of Kabuki, Bardet-Biedl, Cornelia de Lange, and Ehlers-Danlos syndromes.

Sample collection, clinical exome sequencing, whole exome sequencing, and CNV analysis

Sample Collection: Peripheral venous blood samples (2 ml) were collected from patients and their family members. Genomic DNA was extracted from these samples using the MagNA Pure 96 DNA and Viral NA Small Volume Kit.

Sample Preparation: CES was performed using the Roche HyperCap Designed Share (DS) Inherited Disease Panel kit covering 4125 genes, while WES was executed using the Roche Kapa Hyperexome kit. Samples were prepared following the respective kit protocols, which involved capturing exonic regions of interest using targeted probes.

Sequencing: The prepared CES and WES samples were sequenced on the MGI sequencing platform, DNBSEQ-G400. This process generated raw sequencing data in the form of short reads, representing the DNA fragments.

Variant Classification and Analysis: The raw sequencing data (FASTQ files) were uploaded to the SEQ Platform developed by Genomize Inc. The reads were aligned to the human reference genome GRCh37 using the Burrows-Wheeler Aligner [8]. Variant calling was performed using FreeBayes [9]., followed by additional steps such as PCR deduplication and in-del realignment using Genomize's proprietary algorithms. Identified variants were annotated using VEP v102 [10] to provide functional annotations. ACMG pathogenicity classification was conducted using Genomize's proprietary algorithm based on the guideline published by Richards et al. [11].

Copy Number Variation Analysis: CNV analysis was performed using the GATK-gCNV tool, part of the Genome Analysis Toolkit (GATK), version 4.1.8.1. GATK-gCNV utilizes a Bayesian algorithm designed to detect rare CNVs from whole-exome sequencing (WES) data by analyzing sequencing read-depth information. Data Preparation and Preprocessing: Binning was disabled, and all intervals targeted by the whole exome and clinical exome were padded by 250 base pairs (bps). Initially, genomic regions were filtered based on GC content, mappability, and segmental duplication content. Regions with a GC content below 0.1 or above 0.9 were excluded to mitigate GC bias. Additionally, regions with a mappability score below 0.9 or above 1.0 were filtered out to ensure high confidence in read alignment. Genomic intervals with segmental duplication content above 0.5 were excluded to avoid potential false positives. Subsequently, intervals with fewer than 10 read counts in more than 50% of the cohort were filtered out. Extreme read counts falling within the 1st percentile (minimum) and 99th percentile (maximum) across 90% of the cohort were also excluded from further analysis. CNV Filtering and Evaluation: The detected CNVs were first filtered based on their frequency within the same NGS run. CNVs that were observed in multiple patients with different phenotypes were excluded from further analysis. These remaining CNVs were then evaluated using databases such as DGV, ClinVar, and OMIM to assess their pathogenicity. Following this process, pathogenic, likely pathogenic, and variants of uncertain significance that were considered to explain the patient's clinical

findings were reviewed for read depth in IGV (Integrative Genomics Viewer). For clinically relevant CNVs, IGV was used to validate the findings. In cases of homozygous deletions, IGV displayed no reads in the affected region, while in heterozygous deletions, SNPs within the region appeared as homozygous.

Results

The exome sequencing-based CNV analysis yielded an overall diagnostic yield of 2.67% (12/449) in our study. The patients enrolled in the study presented with various indications for genetic testing, including cardiomyopathy, lactic acidosis, spastic paraplegia, neuronal ceroid lipofuscinosis, epidermolysis bullosa, unidentified vision loss with deafness, and preliminary diagnoses of Kabuki, Bardet Biedl, Cornelia de Lange, and Ehlers Danlos syndromes. Clinical exome sequencing (CES) was performed for 5 patients, whole exome sequencing (WES) for 3 patients, and trio-WES for 4 patients. The NGS-based CNV analysis revealed 12 distinct CNVs that correlated with the patients' clinical presentations. These included monoallelic deletions in the *KMT2D*, *EIF5A*, and *LAMP2* genes, as well as biallelic deletions in the *ZFYVE26*, *AP4S1*, *PPT1*, *PDHX*, *BBS9*, *COL7A1*, and *RPGRIP1* genes. Furthermore, a monoallelic duplication was identified in the *HDAC8* gene, and a biallelic duplication was observed in the *PLOD1* gene. Biallelic deletions were seen in the Integrative Genomics Viewer (IGV). Figure 1A, B, C and D display coverage images from some the patients with homozygous deletions alongside controls, as visualized in the IGV.

The sizes of the detected CNVs ranged from approximately 485 base pairs (bp) for the smallest deletion to 228,307 bp for the largest deletion.

Table 1 summarizes the clinical findings of the patients and presents the results of the NGS-based CNV analysis., Table 2 shows the average read depth, 20X and 50X coverage, and the total number of CNVs detected per sample. Figure 2 illustrates the dysmorphologic features of patients 1, 2, and 8.

Family segregation studies utilizing NGS-based CNV analysis were conducted in six patients. The results of the segregation analysis in all six patients were consistent with the inheritance models of the diseases. Specifically,

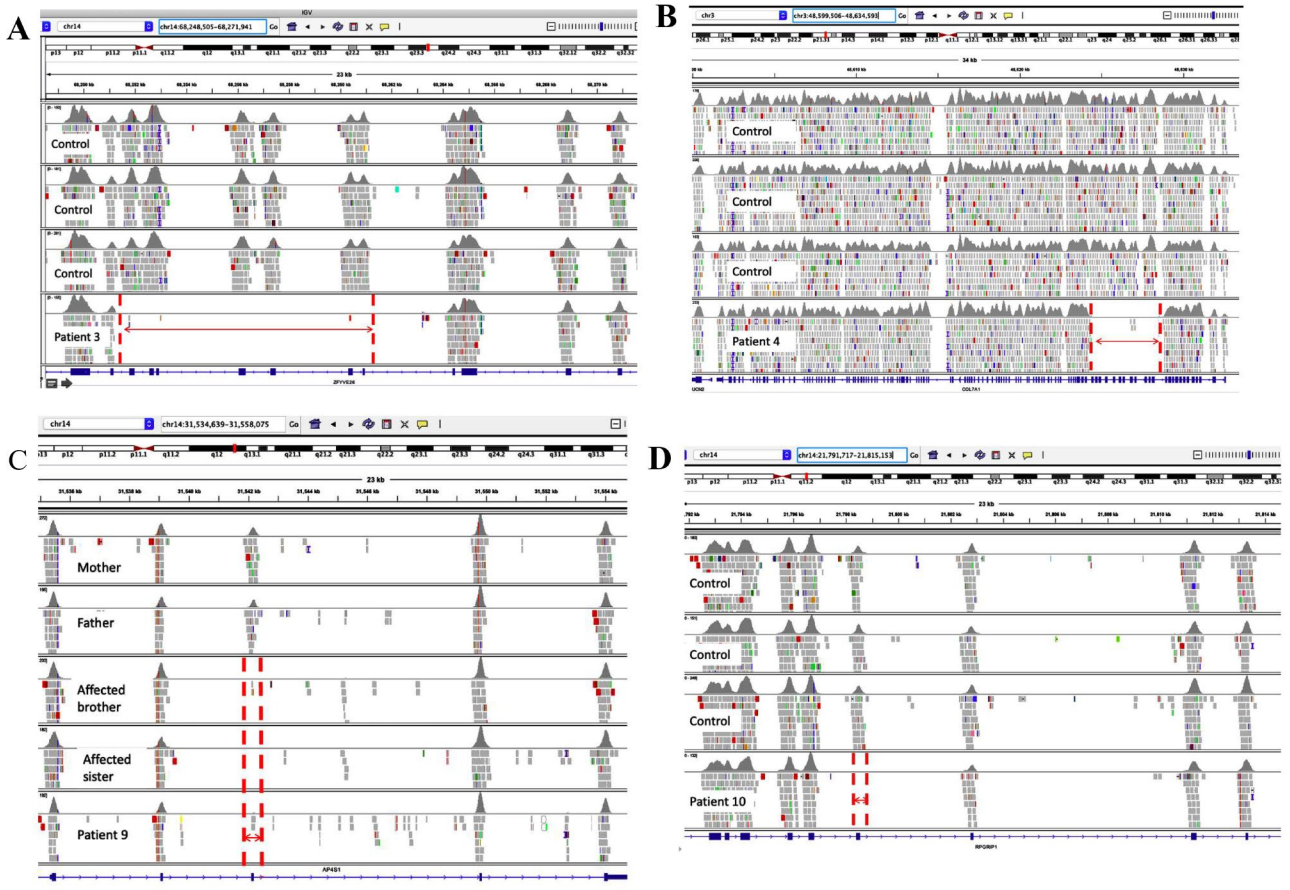

Fig. 1 **A**Biallelic deletion of the Exon 13–19 in the *ZFYVE26* gene in the Patient 3. **B** Biallelic deletion of exons 13 to 24 in the *COL7A1* gene was in Patient 4. **C** Biallelic deletion of the Exon 4 in the *AP4S1* gene in the patient 9 and affected siblings. **D** Biallelic deletion of the Exon 19 in the *RPGRIP1* gene in the Patient 10.

Table 1 Clinical and molecular features of the patients

Patient No	Clinical findings	Parents Consanguinity	Test	CNV- GRCh37 approximate position	CNV type Gene / Exon	Previous data	Diagnosis
1	15-months-old-girl, developmental delay, cleft palate, dysmorphic features	No	CES	chrX: (71681842–71710877)x3 (29036 bp)	Dup *HDAC8* (NM_018486.3) Exon 6–9	Kaiser (2014) Hum Mol Genet 23, 2888 [13]	Cornelia de Lange syndrome 5
2	5-year-old-girl, developmental de lay, dysmorphic features	No	WES	chr12: (49446308–49448006)x1 (1699 bp)	Het, Del *KMT2D* (NM_003482.4) Exon 5–9	NR	Kabuki syndrome 1
3	17-year-old-boy, progressive spasticity	Yes	CES	chr14: (68251517–68261225)x0 (9709 bp)	Hom, Del *ZFYVE26* (NM_015346.4) Exon 13–19	NR	Spastic paraplegia 15, autosomal recessive
4	3-day-old baby, epidermolysis bullosa	Originated from the same village	Trio WES	chr3: (48624264–48628519)x0 (4256 bp)	Hom, Del *COL7A1* (NM_000094.3) Exon 13–24	Taghizadeh M, Our Dermatol Online. 2021 [14]	Epidermolysis bullosa dystrophica
5	6-year-old girl microcephaly, seizures, congenital lactic asidosis	Yes	CES	chr11: (34978672–34982326)x0 (3655 bp)	Hom, Del *PDHX* (NM_003477.2) Exon 4–5	NR	Lacticacidemia due to PDX1 deficiency
6	22-month-old boy rapid psychomotor deterioration, refracter seizures	Yes	CES	chr1 (40557954–40558438)x0 (485 bp)	Hom, Del *PPT1* (NM_000310.3) Exon 2	Simonati (2009) Pediatr Neurol 40, 271 [15]	Ceroid lipofuscinosis, neuronal, 1
7	4-year-old boy soft skin, marked joint laxity, thoracolumbar kyphoscoliosis	Originated from the same village	Trio WES	chr1 (12020452–12027403)x4 (6952 bp)	Hom, Dup *PLOD1* (NM_000302.3) Exon 10–16	Zahed-Cheikh (2017) Indian Pediatr 54, 495 [16]	Ehlers-Danlos syndrome, kyphoscoliotic type
8	4-year-old boy microcephaly, developmental delay dysmorphic features	No	Trio WES	chr17 (7004634–7232940)x1 (228307 bp)	Het, Del *EIF5A*(NM_001970.4) *DLG4*(NM_001365.3) Whole gene	NR	Faundes-Banka syndrome, Intellectual developmental disorder, autosomal dominant 62
9	9-year old boy developmental delay, progressive spasticity optic atrophy	Yes	WES, parental CES	chr14 (31541861–31542481)x0 (621 bp)	Hom, Del *AP4S1* (NM_001128126.3) Exon 4	NR	Spastic paraplegia 52, autosomal recessive

Table 1 (continued)

Patient No	Clinical findings	Parents Consanguinity	Test	CNV- GRCh37 approximate position	CNV type Gene / Exon	Previous data	Diagnosis
10	14-month-old girl congenital sensorineural deafness, nistagmus, decrease in visual acuity	Yes	CES	chr14 (21798167–21798786)x0 (620 bp)	Hom, Del *RPGRIP1* (NM_020366.3) Exon 19	Ellingford (2017) Eur J Hum Genet 25, 719 [2]	Leber congenital amaurosis 6
11	15-year-old boy, hypertrophic cardiomyopathy	No	Trio WES	chrX (119581436–119583242)x0 (1807 bp)	Hemi, Del *LAMP2* (NM_002294.3) Exon 4–5	NR	Danon disease
12	9-year-old girl retinal dystrophy, polydactyly, polyneuropathy, myosite episodes	Yes	WES, parental CES	chr7 (33423019–33428015)x0 (4997 bp)	Hom, Del *BBS9* (NM_198428.2) Exon 18–19	NR	Bardet-Biedl syndrome 9

Abbreviation Hom: homozygous, Het: Heterozygous, Hemi: Hemizygous, Del: deletion, Dup: duplication, NR: not reported

trio-WES was performed for four patients (Patients 4, 7, 8, and 11). For the parents of Patient 12, as well as the parents and affected siblings of Patient 9, CES was carried out. It is worth noting that in two patients (Patient 10 and 12), a dual molecular diagnosis was established, with one diagnosis attributed to a single nucleotide variant and the other diagnosis resulting from a CNV. Below, we present detailed clinical findings for some of the patients. Further information about the remaining patients can be accessed through the supplementary material.

Patient 8

Patient 8 was a 4-year-old boy evaluated due to microcephaly, failure to thrive, developmental delay and dysmorphic features including laterally sparsed eyebrowes, long eyelashes, long and down slanting palpebral fissures, bulbous nose and cupped ears suggesting Kabuki syndrome. Trio WES was performed and there was no disease-causing single nucleotide variation. CNV analysis revealed a 228-kb heterozygous deletion on chromosomal region 17p13.1 including *EIF5A* and *DLG4* genes responsible for Faundes-Banka syndrome and Autosomal dominant intellectual developmental disorder-62, respectively. NGS-based CNV analysis in the parents' WES did not detect this CNV, suggesting its de novo origin.

Patient 10

Patient 10 was a 14-month-old girl born to consanguineous marriage presented with mild developmental delay, congenital sensorineural deafness, decrease in visual acuity, horizontal and rotatuar nistagmus. On physical examination, she had sparse hair and eyebrows, frontal bossing, esotropia in the right eye, depressed nasal bridge, bulbous nasal tip and retromicrognathia. Optic nerves were evaluated as normal and retinal vessels were found thinned in the eye examination. In the CES analysis, there was a homozygous, likely pathogenic variant (c.3679 C>T, p.Arg1227Ter) in the *OTOF* (NM_194248.3) gene (Fig. 3) responsible for autosomal recessive nonsyndromic hearing loss 9 consistent with sensorineural hearing loss. CNV analysis revealed a homozygous, 620 bp deletion including exon 19 of the *RPGRIP1* (NM_020366.3) gene responsible for Cone-rod dystrophy 13 and Leber congenital amaurosis 6.

Patient 12

Patient 12, a 9-year-old girl, presented with retinal dystrophy, postaxial polydactyly, axonal polyneuropathy, and myositis episodes during infections. Whole exome sequencing was performed and CNV analysis revealed a homozygous deletion in the *BBS9* (NM_198428.2), gene specifically including Exon 18–19, which is responsible for Bardet-Biedl Syndrome type 9. Additionally, Patient 12 had a homozygous missense variant (c.487G>A,

Table 2 The average read depth, 20X and 50X coverage, and the total number of CNVs detected per sample

Patient No	Test	Average Read Depth	20X Co-varage %	50X Co-varage %	Total number of CNVs detected
1	CES	80X	98	84	11
2	WES	78X	98	82	85
3	CES	88X	98	88	7
4	WES	118X	99	96	92
5	CES	186X	99	98	4
6	CES	102X	99	92	9
7	WES	117X	99	95	88
8	WES	41X	91	27	81
9	WES	109X	99	95	84
10	CES	56X	96	60	18
11	WES	84X	99	86	114
12	WES	77X	98	80	96

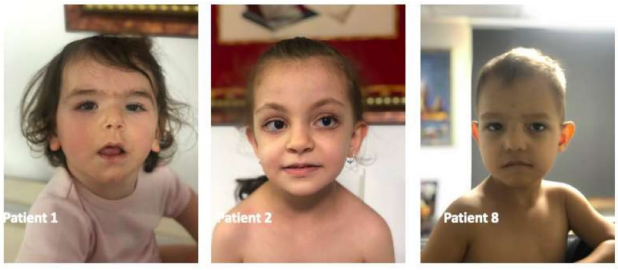

Fig. 2 Dysmorphologic features of patients **1**, **2**, and **8**. (Patient 1 has arched eyebrows, synophrys, long eyelashes, depressed nasal bridge, long filtrum and pointed chin. Patient 2 has arched eyebrows, long palpebral fissures, eversion of lateral third of lower eyelids, left iris coloboma, large prominent ears, thin upper lip vermillion. Patient 8 has laterally sparsed eyebrowes, long eyelashes, long and down-slanting palpebral fissures, bulbous nose and cupped ears)

p.Gly163Ser) classified as "likely pathogenic" in the *HADHB* (NM_000183.3) gene which is responsible for Mitochondrial Trifunctional Protein Deficiency 1. This variant explains the patient's myositis episodes and axonal polyneuropathy. Clinical exome sequencing were performed for segregation analysis and both of the parents were heterozygous for the deletion of exon 18–19 in the *BBS9* gene and the single nucleotide variant in the *HADHB* gene.

Discussion

In this study, we performed exome-based CNV screening via a read-depth analysis of exome sequencing data from a large patient cohort with various phenotypes. While prior studies have demonstrated the efficacy of exome-based CNV analysis in identifying genomic CNVs, it's noteworthy that the specificity of the technique is contingent on the employed CNV detection algorithms [2–4, 17]. For our cohort, the read-depth CNV analysis yielded an overall diagnostic improvement of 2.67%, spanning a diverse array of genetic disorders. This finding is in concordance with the diagnostic yield obtained through NGS-based CNV analysis in heterogeneous patient cohorts including 2418 and 2603 patients, 2% and 1,7% respectively [3, 18].

Disease-specific studies have reported varying diagnostic CNV detection rates. For instance, an 18.92% diagnostic rate was achieved using exome-based CNV analysis in a retrospective study involving 74 families afflicted with neurodevelopmental disorders [19]. In our study, patient 1, under evaluation for neurodevelopmental delay and dysmorphic features, was found to possess a duplication in exons 6–9 of the *HDAC8* gene, previously documented by Kaiser et al. [13], leading to a diagnosis of Cornelia de Lange Syndrome Type 5. In patient 2, initially diagnosed with Kabuki syndrome, our NGS-based CNV analysis identified a novel deletion in the *KMT2D* gene. In a compelling case, Patient 8, who exhibited neurodevelopmental delay and dysmorphic features akin to those seen in Kabuki syndrome, was found to carry a de novo, heterozygous, total deletion of the *EIF5A* and *DLG4* genes. These genes are known to be associated with Faundes-Banka syndrome and Intellectual Developmental Disorder, autosomal dominant 62, respectively. Notably, Faundes-Banka syndrome, first described by Faundes et al. in 2021 [20], is a rare disorder that arises as a result of loss-of-function mutations in the *EIF5A* gene. Up until now, only seven patients with this syndrome have been reported. Intriguingly, among these patients, three initially presented with clinical features that mimicked Kabuki syndrome, reminiscent of our patient's case [20]. By identifying this unusual genetic alteration, our study contributes to broadening both the clinical and molecular spectrum associated with Faundes-Banka syndrome.

A comprehensive study spearheaded by Ceyhan-Birsoy et al., meticulously evaluated over 1400 patients diagnosed with cardiomyopathy. Their efforts led to the discovery of clinically significant CNVs in approximately 0.63% of the examined population [21]. Notably, one of the affected genes identified among the detected deletions was the *LAMP2* gene responsible for Danon disease, mirroring the genetic findings observed in patient 11 from our cohort. Evidence from previous research has indicated the presence of 12 substantial deletions in the *LAMP2* gene [22]. While the work by Ceyhan-Birsoy et al. suggests that the majority of cardiomyopathy cases may not be precipitated by small-scale CNVs, it is essential to not rule out this possibility, especially in instances where the fundamental molecular cause remains unidentified.

To our current knowledge, more than 80 genes have been implicated in the etiology of hereditary spastic paraplegias (HSPs) [23]. A recent comprehensive review conducted by Galatolo et al. [24], which synthesized findings

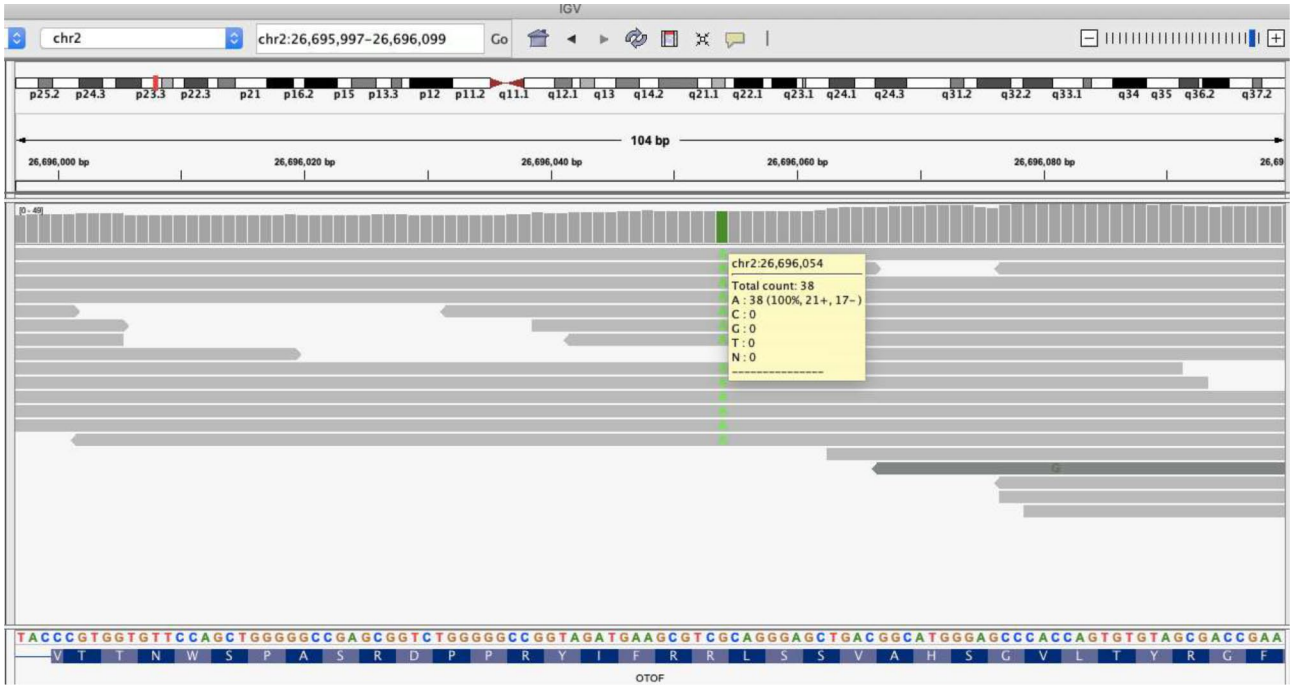

Fig. 3 Homozygous c.3679 C > T, (p.Arg1227Ter) variant in the *OTOF* (NM_194248.3) gene

from 19 WES studies, revealed that the rate of molecular etiology detection in HSPs stands at an aggregate of 49.7%. Despite these strides in understanding the genetic basis of HSPs, the role of CNVs in the onset and progression of these disorders remains insufficiently elucidated. Our study further contributes to this evolving field by identifying homozygous deletions in the *ZFYVE26* and *AP4S1* genes in two cases preliminarily diagnosed with HSP. These genes are known to be causative for Spastic Paraplegia 15 and Spastic Paraplegia 52, respectively. While the *ZFYVE26* gene has had two major deletions previously reported in the Human Gene Mutation Database [22], our study has the distinction of reporting a substantial deletion in the *AP4S1* gene for the first time.

In a study where 500 patients with inherited retinal diseases were evaluated, the molecular etiology of 8.8% of patients was successfully identified through the application of NGS-based CNV analysis [25]. This method was particularly effective in diagnosing a variety of ocular disorders. In one noteworthy case (Patient 10), a deletion in exon 19 of the *RPGRIP1* gene associated with Cone-rod dystrophy 13 and Leber congenital amaurosis 6 was detected. The observed deletion has been previously documented and is one among 14 major deletions reported in the *RPGRIP1* gene, according to the Human Gene Mutation Database [2, 22]. In another patient (Patient 12), who was evaluated for retinal dystrophy and postaxial polydactyly, exome-based CNV analysis was instrumental in identifying a homozygous deletion of exons 18 and 19 in the *BBS9* gene which is responsible for Bardet-Biedl Syndrome type 9. Although deletions in exons 18 and 19 of the BBS9 gene had not been previously reported, the gene itself is not novel to CNV research, with a total of 11 different large deletions having been documented to date [22].

Epidermolysis bullosa (EB) is a rare hereditary genetic disorder characterized by fragile skin and blister formation, associated with mutations in 16 different genes, highlighting its genetic heterogeneity [26]. One of these genes, *COL7A1*, is implicated in dystrophic EB, a particularly severe form of the condition. In cases of dystrophic EB, pathogenic variants are detectable through sequence analysis in approximately 95% of cases. However, for the autosomal recessive form of dystrophic EB, less than 2% of cases present pathogenic variants identifiable through deletion/duplication analysis [27]. During the analysis of Patient 4 in this study, NGS-based CNV analysis proved invaluable. This method identified a deletion encompassing exons 13–24 of the COL7A1 gene, a finding previously reported by Taghizadeh et al. [14]. The identification of this molecular etiology is not just an academic triumph; it has immediate practical implications. Not only does it allow for improved genetic counseling, enhancing our understanding of inheritance patterns and risks, but it also paves the way for novel therapeutic interventions. For instance, a local gene therapy called Beremagene Geperpave has recently demonstrated effectiveness in treating dystrophic EB [28] and has received FDA approval [29]. Therefore, understanding the specific genetic makeup of a patient's disease can significantly

impact the treatment strategies available, potentially leading to improved patient outcomes.

Establishing the molecular cause of genetically diverse disease groups typically involves a variety of testing methodologies. These often include targeted NGS panels, clinical exome sequencing or whole exome sequencing. These methods are utilized to explore a broad spectrum of genetic variability, from single nucleotide changes to large-scale chromosomal rearrangements. One of the significant advantages of NGS-based CNV analysis is its ability to concurrently evaluate both single nucleotide variations and CNVs across multiple genes. In the present study, we were able to establish a dual molecular diagnosis for two patients (Patient 10 and 12). Each diagnosis involved distinct molecular events; one was attributed to a single nucleotide variant and the other was a consequence of a CNV. This makes it an especially powerful tool in studying genetically heterogeneous conditions, where pathogenic variations can occur in several different genes.

In the context of this study, NGS-based CNV analysis was used to evaluate patients with genetically diverse conditions such as Ehlers-Danlos syndrome, neuronal ceroid lipofuscinosis, congenital lactic acidosis, hereditary spastic paraplegia, and epidermolysis bullosa. Each of these conditions exhibits significant genetic heterogeneity, underscoring the utility of a testing methodology capable of broad and simultaneous analysis of multiple genes. The use of NGS-based CNV analysis in this setting also served to overcome the limitations of MLPA. It is a commonly used technique for the detection of CNVs, but it is restricted to the evaluation of specific targeted regions using designated probes. Furthermore, it incurs additional costs due to the need for separate probe sets for each gene or region of interest [6]. NGS-based CNV analysis, in contrast, provides a more holistic view of the genomic landscape, enabling the detection of a wider range of potential pathogenic variants and paving the way for comprehensive genetic diagnoses.

One of the limitations of our study is that we did not use a secondary method to confirm the CNVs detected by NGS-based CNV analysis, nor did we employ multiple CNV callers, which could enhance sensitivity and reduce inconsistencies. Despite this, all CNVs identified were consistent with the clinical findings of each patient, reinforcing the validity of our results. For instance, single-exon deletions explaining the clinical presentations were detected in Patients 6, 9, and 10. In Patient 6, the palmitoyl-protein thioesterase activity was measured as "0," aligning with the clinical diagnosis, and in Patient 9, exome-based family segregation analysis confirmed the expected inheritance pattern for the identified CNV.

Although GATK-gCNV is a reliable tool for detecting CNVs across multiple exons, its sensitivity decreases for single-exon deletions. Despite this, single-exon deletions were successfully identified in our study, as highlighted in Patients 6, 9, and 10. The NGS-based CNV analysis method employed in this study demonstrates a recall rate of 97% for rare coding CNVs (with a site frequency of ≤1%) compared to those identified by microarray analysis, and a robust recall rate of 95% for rare coding CNVs identified by genome sequencing, with a resolution exceeding two exons [12].

Our findings underscore the fact that exome-based CNV analysis significantly enhances molecular diagnostic success across a range of disease groups. Consistent with these findings, a recent statement by the American College of Medical Genetics and Genomics (ACMG) emphasizes the critical role structural variants play in understanding the genetic basis of diseases. The statement also highlights the inherent challenges associated with detecting these variants due to their complex nature. While it acknowledges the limitations of traditional cytogenetic methods, it accentuates the superior resolution and comprehensive genome coverage offered by NGS-based approaches in detecting structural variants [30].

In conclusion, our results underscore the effectiveness of NGS-based CNV analysis in determining the molecular etiology of Mendelian diseases. The precise detection of CNVs through this method showcases its potential as a robust tool in genetic diagnostics. Therefore, our study underscores the necessity of integrating NGS-based CNV analysis into regular genetic testing, which could potentially result in improved clinical outcomes and individualized management for people with Mendelian disorders.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-024-02015-1.

Supplementary Material 1

Acknowledgements

We are grateful to the patients and their families who participated in this study.

Author contributions

TA, EAD, EI: Conception and design, acquisition of data, analysis and interpretation of data, writing original draft. MK, SK: Acquisition of data. AA, AD: Analysis and interpretation of data. TA, FO, OC: Analysis and interpretation of data, writing original draft, manuscript revising.EAD accepts full responsibility for the work and/or conduct of the study, had access to the data, controlled the decision to publish, and acts as the study's guarantor.

Funding

No specific funding.

Data availability

The datasets generated and/or analysed during the current study are available in ClinVar with accession numbers SCV005061998, SCV005061999, SCV005062000, SCV005062001, SCV005062002, SCV005062003, SCV005062004 and can be accessed through this link: https://www.ncbi.nlm.nih.gov/clinvar/.

Declarations

Ethics approval and consent to participate

The present study was approved by the Ethics Committee of Ege University Faculty of Medicine (Reference No: 23-7T/27). Written informed consent to participate was obtained from all participants and their parents/legal guardians.

Consent for publication

Written informed consent for publication of their clinical details and clinical images was obtained from all participants and their parents/legal guardians.

Competing interests

The authors declare no competing interests.

Received: 26 June 2024 / Accepted: 19 September 2024
Published online: 30 September 2024

References

1. Shaikh TH, et al. Copy Number Variation disorders. Curr Genet Med Rep. 2017;5(4):183–90.
2. Ellingford JM, Campbell C, Barton S, Bhaskar S, Gupta S, Taylor RL, et al. Validation of copy number variation analysis for next-generation sequencing diagnostics. Eur J Hum Genet. 2017;25(6):719–24.
3. Pfundt R, Del Rosario M, Vissers LELM, Kwint MP, Janssen IM, de Leeuw N, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. Genet Med. 2017;19(6):667–75.
4. Royer-Bertrand B, Cisarova K, Niel-Butschi F, Mittaz-Crettol L, Fodstad H, Superti-Furga A, et al. CNV Detection from Exome Sequencing Data in Routine Diagnostics of Rare Genetic disorders: opportunities and limitations. Genes (Basel). 2021;12(9):1427.
5. Manning M, Hudgins L, Palmer CG, Barch MJ, Nussbaum RL, Hoyme HE, et al. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. Genet Med. 2010;12(11):742–5.
6. Stuppia L, Antonucci I, Palka G, Gatta V, Palka C, De Angelis MV, et al. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. Int J Mol Sci. 2012;13(3):3245–76.
7. Kadalayil L, Rafiq S, Rose-Zerilli MJ, Pengelly RJ, Parker H, Oscier D, et al. Exome sequence read depth methods for identifying copy number changes. Brief Bioinform. 2015;16(3):380–92.
8. Li H, Durbin R, et al. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
9. Garrison E, Marth G et al. Haplotype-based variant detection from short-read sequencing. arXiv Preprint arXiv:12073907 2012.
10. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl variant effect predictor. Genome Biol. 2016;17(1):122.
11. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, ACMG Laboratory Quality Assurance Committee, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–24.
12. Babadi M, Fu JM, Lee SK, Smirnov AN, Gauthier LD, Walker M, et al. GATK-gCNV enables the discovery of rare copy number variants from exome sequencing data. Nat Genet. 2023;55(9):1589–97.
13. Kaiser FJ, Ansari M, Braunholz D, Concepción Gil-Rodríguez M, Decroos C, Wilde JJ, et al. Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia De Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. Hum Mol Genet. 2014;23(11):2888–900.
14. Taghizadeh M, Mansoori Derakhshan S, Shekari Khaniani M, Eshaghkhani Y, Golchehre Z, Taheri SR, et al. Identification of multi-exon deletion in the COL7A1 gene underlying dystrophic epidermolysis bullosa by whole-exome sequencing. Our Dermatol Online. 2021;12(4):412–6.
15. Simonati A, Tessa A, Bernardina BD, Biancheri R, Veneselli E, Tozzi G, et al. Variant late infantile neuronal ceroid lipofuscinosis because of CLN1 mutations. Pediatr Neurol. 2009;40(4):271–6.
16. Zahed-Cheikh M, Tosello B, Coze S, Gire C, et al. Kyphoscolitic type of Ehlers-Danlos syndrome with prenatal stroke. Indian Pediatr. 2017;54(6):495–7.
17. Moreno-Cabrera JM, Del Valle J, Castellanos E, Feliubadaló L, Pineda M, Brunet J, et al. Evaluation of CNV detection tools for NGS panel data in genetic diagnostics. Eur J Hum Genet. 2020;28(12):1645–55.
18. Testard Q, Vanhoye X, Yauy K, Naud ME, Vieville G, Rousseau F, et al. Exome sequencing as a first-tier test for copy number variant detection: retrospective evaluation and prospective screening in 2418 cases. J Med Genet. 2022;59(12):1234–40.
19. Zhai Y, Zhang Z, Shi P, Martin DM, Kong X, et al. Incorporation of exome-based CNV analysis makes trio-WES a more powerful tool for clinical diagnosis in neurodevelopmental disorders: a retrospective study. Hum Mutat. 2021;42(8):990–1004.
20. Faundes V, Jennings MD, Crilly S, Legraie S, Withers SE, Cuvertino S, et al. Impaired eIF5A function causes a mendelian disorder that is partially rescued in model systems by spermidine. Nat Commun. 2021;12(1):833.
21. Ceyhan-Birsoy O, Pugh TJ, Bowser MJ, Hynes E, Frisella AL, Mahanta LM, et al. Next generation sequencing-based copy number analysis reveals low prevalence of deletions and duplications in 46 genes associated with genetic cardiomyopathies. Mol Genet Genomic Med. 2015;4(2):143–51.
22. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, et al. The human gene mutation database (HGMD®): 2003 Update. Hum Mutat. 2003;21:577–81.
23. Hedera P et al. Hereditary Spastic Paraplegia Overview. In: Adam MP, Mirzaa GM, Pagon RA, editors. GeneReviews® [Internet]. University of Washington, Seattle; 1993–2023.
24. Galatolo D, Trovato R, Scarlatti A, Rossi S, Natale G, De Michele G et al. Power of NGS-based tests in HSP diagnosis: analysis of massively parallel sequencing in clinical practice. Neurogenetics 2023 May 3.
25. Zampaglione E, Kinde B, Place EM, Navarro-Gomez D, Maher M, Jamshidi F, et al. Copy-number variation contributes 9% of pathogenicity in the inherited retinal degenerations. Genet Med. 2020;22(6):1079–87.
26. Bardhan A, Bruckner-Tuderman L, Chapple ILC, et al. Epidermolysis Bullosa. Nat Rev Dis Primers. 2020;6:78.
27. Pfendner EG, Lucky AW et al. Dystrophic Epidermolysis Bullosa. In: Adam MP, Mirzaa GM, Pagon RA, editors. GeneReviews® [Internet]. University of Washington, Seattle; 1993–2023.
28. Guide SV, Gonzalez ME, Bağcı IS, Agostini B, Chen H, Feeney G, et al. Trial of Beremagene Geperpavec (B-VEC) for Dystrophic Epidermolysis Bullosa. N Engl J Med. 2022;387(24):2211–9.
29. Food and Drug Administration Web site. https://www.fda.gov/vaccines-blood-biologics/vyjuvek [Accessed June 23, 2023].
30. Raca G, Astbury C, Behlmann A, De Castro MJ, Hickey SE, Karaca E, ACMG Laboratory Quality Assurance Committee, et al. Points to consider in the detection of germline structural variants using next-generation sequencing: a statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2023;25(2):100316.

Publisher's note