Parallel Tests of Whole Exome Sequencing and Copy Number Variant Sequencing Increase the Diagnosis Yields of Rare Pediatric Disorders

Xuyun Hu [1,2,3], Ruolan Guo [1,2,3], Jun Guo [1,2,3], Zhan Qi [1,2,3], Wei Li [1,2,3] and Chanjuan Hao [1,2,3]*

[1] Beijing Key Laboratory for Genetics of Birth Defects, Beijing Pediatric Research Institute, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [2] MOE Key Laboratory of Major Diseases in Children, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China, [3] Henan Key Laboratory of Pediatric Inherited & Metabolic Diseases, Henan Children's Hospital, Zhengzhou Hospital of Beijing Children's Hospital, Zhengzhou, China

Background: Both whole exome sequencing and copy number variants sequencing were applied to identify the genetic cause of rare pediatric disorders. In our study, we aimed to investigate the diagnostic yield of parallel tests of trio whole exome sequencing and copy number variants sequencing and its clinical utility.

Methods: After collecting detailed clinical information, a total of 60 patients were referred to parallel tests of whole exome sequencing and copy number variants sequencing, which used shared initial libraries.

Results: 26 pathogenic or likely pathogenic single nucleotide variants and 11 copy number variants were identified in 32 patients. 65.4% (17/26) of the SNVs were novel. The overall diagnosis rate was 53.3%. For the patients with positive results, 22 (36.7%) patients were diagnosed by whole exome sequencing and 10 (16.7%) patients were diagnosed by copy number variants sequencing. We also reviewed clinical impact on selected cases.

Conclusion: We adopted an approach by performing parallel tests of trio whole exome sequencing and copy number variants sequencing with shared initial libraries. This strategy is relatively efficient and cost-effective for the diagnosis of rare pediatric disorders with high heterogeneity.

Keywords: whole exome sequencing, copy number variants sequencing, pediatric disorders, cost-effective, clinical utility

INTRODUCTION

High genetic heterogeneity in pediatric disorders has been an obstacle to phenotype-based diagnostic testing (Hu et al., 2018). Compared to conventional genetic tests on single-gene scale, tests on genomic scale are more likely to produce higher diagnostic yields and shorter turn-around times (Vissers et al., 2017). Generally, these genomic tests include whole exome sequencing (WES)

and chromosomal microarray analysis (CMA). WES has been ordered increasingly in clinical molecular diagnosis laboratories as a powerful tool for rare Mendelian disorders, especially for genetically heterogeneous disorders such as intellectual developmental disorders and multiple congenital anomalies (Bowling et al., 2017; Han and Lee, 2020). According to previous researches, the overall diagnostic yields among different cohorts were ~25% (Yang et al., 2013, 2014; Lee et al., 2014). Recently, multiple researches on clinical utility and cost of WES have provided evidence endorsing it as a first-tier test for children with suspected monogenic disorders (Nguyen and Charlebois, 2015; Monroe et al., 2016; Stark et al., 2016; Hu et al., 2018). CMA has been used as a first-tier clinical diagnostic test in patients with developmental delay (DD)/intellectual disability (ID), autism spectrum disorders (ASD), and multiple congenital anomalies (MCA) since 2010 (Manning et al., 2010; Miller et al., 2010); and copy number variants (CNVs) detected by CMA explained the pathogenesis for over 10% of these cases (Sanmann et al., 2015; Homma et al., 2018; Jang et al., 2019).

Recently, WES or whole genome sequencing (WGS) data were also used to call CNVs through the development of various algorithms and software programs that utilize read-depth information as the main strategy (Wang et al., 2016; Yao et al., 2017). WES-data-based CNV calling is usually performed as a supplement of routine WES test and the accuracy is affected by capture area and amplification efficiency (Rajagopalan et al., 2020; Sun et al., 2020). On the other hand, low-coverage WGS or CNV-seq were widely used in spontaneous miscarriage and prenatal cases, but only a few pediatric CNV-seq testing studies were reported. Most of the studies focused only on neurological disorders (Gao et al., 2019; Jiao et al., 2019). In our study, we performed parallel tests of trio-WES and CNV-seq for 60 children from different clinical departments. Our results showed this parallel testing strategy significantly increased diagnostic yields and lightened the burden of physicians in selecting optimal test.

MATERIALS AND METHODS

Patients

Patients in this research were initially referred to the Beijing Children's Hospital from September 2018 to September 2019, with suspected Mendelian disorders. These patients had DD/ID, ASD, or MCA, and both SNVs and CNVs were highly suspected for their causes of disorders. Next-generation sequencing (NGS) was primarily ordered by the patient's physician and performed by the Laboratory for Genetics of Birth Defects at Beijing Children's Hospital. Written informed consent was obtained from the individuals and their legal guardians for the publication. This study was approved by the Beijing Children's Hospital institutional review board. After recruitment, a total of 60 patients were referred to parallel tests of WES and

CNV-seq. The ages of the patients ranged from 1 month to 12 years, and the median age was 1 year. The male-to-female ratio in this study was 1.31:1 (34/26). For the referring reasons, 25 of the patients had multiple congenital anomalies. Ten had DD/ID. Six had the combination of DD/ID and multiple congenital anomalies. Four had autistic traits. The remaining 15 patients had other phenotype, such as congenital heart disease, short stature, recurrent infections, etc. (**Supplemental Table 1**).

Whole Exome Sequencing

DNA was isolated from peripheral blood samples obtained from the proband and parents using the Gentra Puregene Blood Kit (QIAGEN, Hilden, Germany). 200-ng genomic DNA of each individual was sheared by Biorupter (Diagenode, Liège, Belgium) to acquire 150~200-bp fragments. The ends of DNA fragment were repaired, and Illumina Adaptors were added (Fast Library Prep Kit, iGeneTech, Beijing, China). After sequencing library was constructed, the whole exome was captured with AIExome Enrichment Kit V1 (iGeneTech, Beijing, China) and sequenced on Illumina NovaSeq 6000 (Illumina, San Diego, USA) with 150 base paired-end reads. Raw reads were filtered to remove low-quality reads using FastQC. Clean reads were mapped to the reference genome GRCh37. Quality control (QC) information included: average read length of $>100\times$, accurate mapping rate of $>98\%$, bases capture rate of $>55\%$, $20\times$ mean depth coverage rate of $>96\%$, duplication rate of $<25\%$, and accurate mapping rate of $<96\%$. Single nucleotide variants (SNVs) were annotated and filtered by TGex (tgex.genecards.org). Variants with a frequency over 1% in the databases of gnomAD, ESP or 1000G were excluded. Variants that lacked segregation in family members were also filtered. Variants were classified following the American College of Medical Genetics and Genomics and the Association for Molecular Pathology interpretation standards and guidelines (Richards et al., 2015).

Copy Number Variant Calling

To identify large copy number variants (CNVs), part of the library without capture was sequenced directly onto Illumina NovaSeq 6000, and each sample yielded one giga base (Gb) raw data (QC: average read length: $>0.3\times$ WGS). An in-house pipeline was applied to map and call CNVs based on the software CNV-seq (Xie and Tammi, 2009; Hu et al., 2019). Clean reads were mapped to the reference genome GRCh37. CNVs called from parental WGS data were used as controls. CNVs reported in multiple peer-reviewed publications or annotated in curated databases as benign or likely benign, CNVs observed frequently in the general population and CNVs containing no genes were filtered. Database of Genomic Variants, DECIPHER database, ClinVar, OMIM, and ClinGen were used for interpretation and classification of the clinical significance of candidate CNVs according to previously reported guidelines (Kearney et al., 2011; Riggs et al., 2019). WES data was also used for CNVs calling, and the samples of the same batch were used as controls. Putative pathogenic or likely pathogenic CNVs identified from low-pass WGS were validated using WES reads depth data.

Abbreviations: WES, whole exome sequencing; CNV-seq, copy number variants sequencing; SNVs, single nucleotide variants; CNVs, copy number variants; CMA, chromosomal microarray analysis; DD, developmental delay; ID, intellectual disability; ASD, autism spectrum disorders; MCA, multiple congenital anomalies; WGS, whole genome sequencing; NGS, next-generation sequencing; TAT, turn-around time; QC, quality control.

TABLE 1 | Diagnosis rates of patients with different ordering phenotype categories.

Phenotype	All cases	WES	CNV-seq	All positive	Positive rate (%)
MCA	25	10	3	13	52.0
DD/ID	5	3	0	3	60.0
DD/ID+MCA	12	5	5	10	83.3
ASD	4	0	1	1	25.0
Others	14	4	1	5	42.9
Total	60	22 (36.73%)	10 (16.7%)	32	53.3

ASD, autism spectrum disorders; DD, developmental delay; ID, intellectual disability; MCA, multiple congenital anomalies.

RESULTS

Molecular Diagnosis Rates

We identified 37 pathogenic or likely pathogenic variants in 32 patients. The overall diagnosis rate was 53.3%. For the patients with positive results, 22 (36.7%) individuals were diagnosed by WES (**Table 1**) and 10 (16.7%) individuals were diagnosed by CNV-seq (**Table 2**). For patients with MCA and/or DD/ID, the diagnosis rates were all higher than 50%. For patients with autistic traits, only one individual was diagnosed (**Table 1**).

Characterization of Variants

For the 37 pathogenic or likely pathogenic variants, 26 were SNVs and 11 were CNVs (**Tables 2**, **3**). Twenty-six SNVs were located in 21 different genes, and most of them (17/26, 65.4%) were in autosomal dominant diseases genes. Six variants were located in three autosomal recessive diseases genes, and three variants were located in X-linked recessive diseases genes. 65.4% (17/26) SNVs were not reported in previous literature or public population databases. Thirteen variants were identified to be *de novo*, six variants were compound heterozygous inherited from the parents, five were inherited from parents with similar phenotypes, and two X-linked recessive genes variants inherited from maternal carriers. Mutation in only two genes, *ELN* and *FBN1*, occurred in more than one patient, indicating the high heterogeneity of pediatric disorders (**Table 2**). For 11 CNVs identified by CNV-seq, the sizes ranged from 1.41 to 43.68 Mb. All these variants were also called from WES data to eliminate false positive variants. All these variants were *de novo* (**Table 3**). Patient 23 had a distal deletion and distal duplication, which was potentially inherited from a parent with balanced translocation. Further karyotype test may help to verify this hypothesis.

Turn-Around Time (TAT) and Cost Analysis

We aimed to decrease the difficulty for physicians to select optimal test between WES and CNV-seq, and to shorten TAT, as well, by parallel testing. We tracked the TAT of our parallel test for these patients; it ranged from 38 days to 125 days. The median TAT is 72 days with 69.2% of patients received reports within 80 days. The raw data were usually obtained within 20 days. Initial test results were generated 10–20 days afterwards and were delivered to ordering physicians to further check potential phenotype. Typically, negative cases cost more days because

multiple rounds of communication with ordering physicians would be guaranteed, and a second analysis by another geneticist was performed before the formal reports were issued.

The direct cost of running a parallel trio-WES and CNV-seq is about $600 US dollars ($200 per person), including library construction (~$100 USD), WES (~$80 USD, 100×) and low-coverage WGS (~$10 USD, 0.3×), and Sanger validation (~$10 USD). These results showed that our parallel test strategy was affordable and additional CNV-seq did not increase the prime cost since the generation of one Gb raw data usually costed <$20 USD.

Clinical Impact of Genetic Diagnoses

Over half of the patients ended their diagnosis odysseys after parallel tests. The positive diagnostic results affected clinical management including appropriate genetic counseling, other systemic evaluation, and change of treatment. We provided examples of impacts on patients' clinical managements, below.

Case 4 was a 12-month-old boy. He was referred to Clinic of Development & Behavioral Pediatrics for Short Stature and DD. A *de novo SMARCA2* mutation was identified and was therefore diagnosed as Nicolaides-Baraitser syndrome (Zhang P. et al., 2019). This syndrome was less recognizable and always misdiagnosed as Coffin-Siris syndrome, Williams syndrome, etc. Once molecular diagnosis was confirmed, this patient was referred to neurologist for seizure evaluation. Ophthalmological and audiological examinations were also ordered.

Case 40 was a 5-month-old boy who had developmental and growth delays. Mutations in two genes were identified. DD was caused by *SETD5* mutation inherited from his mother with intellectual disability and growth delay was caused by *ACAN* mutation inherited from his father with short stature. Referral to an early intervention program is recommended for access to occupational, physical, speech, and feeding therapy. Growth pattern should be monitored and growth hormone therapy can improve height increase (Hu et al., 2017).

Case 29 was a 15-year-old girl. She was first referred to Department of Respiration for Pulmonary Lesions. Further evaluation revealed short stature, ataxia, tooth agenesis, depigmentation/hyperpigmentation of skin and absence of secondary sex characteristics. Parallel tests identified a 3.1-Mb *de novo* interstitial deletion of the 14q13.2q21.1 region encompassing 17 OMIM genes (Hu et al., 2019). In these genes, *NKX2*-1 deletion is responsible for choreoathetosis, hypothyroidism, and neonatal respiratory distress and haploinsufficiency of *PAX9* causes oligodontia phenotype (Das et al., 2002; Santen et al., 2012; Hayashi et al., 2015). For this patient, thyroid function testing should be performed annually, and when hypothyroidism is discovered, thyroid hormone replacement therapy should be initiated. For choreoathetosis, Tetrabenazine and Levodopa therapy have been reported to effectively reduce chorea (Setter et al., 2009; Rosati et al., 2015).

Cases 1, 13, and 59 all had 22q11.2 deletion syndrome. This was the only recurrent CNV disorder in our cohort. Case 1 was initially considered as DiGeorge syndrome for the characteristic facial features. Case 13 was referred only for autistic behaviors, and Case 59 was referred for congenital heart

TABLE 2 | Patients with pathogenic and likely pathogenic variants identified by whole exome sequencing.

Case no.	Sex	Age (y)	Categories	Gene	Variant	Origin	Inheritance	Reported*/novel	Diagnosed disorders
4	Male	1	DD/ID+MCA	SMARCA2	c.3593T>G/ p.Val1198Gly	de novo	AD	Reported	Nicolaides-Baraitser syndrome
6	Female	0.33	MCA	RNU4ATAC	c.196-570C>T; c.196-605C>T	in-trans	AR	Reported; reported	Roifman syndrome
8	Male	6	Others	PPP2R5D	c.1492C>T/ p.Arg498*	de novo	AD	Novel	Mental retardation, autosomal dominant 35
10	Male	5	Others	ELN	c.1621C>T/ p.Arg541*	de novo	AD	Novel	Supravalvar aortic stenosis
14	Male	1	Others	DPH1	c.471delC/ p.Arg158fs; c.397T>C/ p.Tyr133His	in-trans	AR	Novel; novel	Developmental delay with short stature, dysmorphic features, and sparse hair
16	Male	0.67	DD/ID	SLC22A5	c.1130T>C/ p.Phe377Ser; c.1400C>G/p.Ser467Cys	in-trans	AR	Novel; reported	Carnitine deficiency, systemic primary
22	Female	3	MCA	MED12	c.6340C>T/ p.Gln2114*	de novo	XLR	Novel	Ohdo syndrome
24	Male	5	MCA	CHAF1A	c.1750C>T / p.Arg584*	de novo	AD	Novel	novel disorder
27	Male	0.08	MCA	COL1A1	c.4100C>A/ p.Thr1367Asn	Maternal	AD	Novel	Osteogenesis imperfecta
31	Female	3	DD/ID+MCA	NSD1	c.6367_6376delAGTTG TGGGG/ p.Ser2123fs	de novo	AD	Novel	Sotos syndrome 1
33	Female	6	DD/ID	KAT6B	c.3663_3664+2delAAGT /p.Asn1222fs	de novo	AD	Novel	SBBYSS/Genitopatellar syndrome
35	Male	4	MCA	MID1	c.1863_1879dupGAACTC CATCCACCTCT/ p.Tyr627fs	Maternal	XLR	Novel	Opitz GBBB syndrome, type I
38	Male	0.17	MCA	FGFR2	c.869G>T/ p.Trp290Leu	de novo	AD	Reported	Crouzon syndrome
40	Male	0.42	DD/ID	ACAN SETD5	c.7000C>T/ p.Gln 2334*; c.1357C>T/ p.Gln453*	Paternal maternal	AD AD	Novel; novel	Short stature and advanced bone age Mental retardation, autosomal dominant 23
41	Female	1	Others	ELN	c.800-1G>T	Maternal	AD	Reported	Supravalvar aortic stenosis
42	Female	1	MCA	FBN1	c.7180C>T/ p.Arg2394*	Paternal	AD	Reported	Marfan syndrome
46	Male	3	DD/ID+MCA	ATRX	c.6254G>A/ p.Arg2085His	Maternal	XLR	Reported	Alpha-thalassemia/ mental retardation syndrome
47	Male	1	DD/ID+MCA	ANKRD11	c.1681G>T/ p.Glu561*	de novo	AD	Novel	KBG syndrome
49	Male	0.42	MCA	PIK3CA	c.2740G>A/ p.Gly914Arg	de novo	AD	Reported	Megalencephaly-capillary malformation
50	Male	12	MCA	FBN1	c.1995C>A/ p.Tyr665*	de novo	AD	Novel	Marfan syndrome
54	Male	0.25	DD/ID+MCA	SOX9	c.340G>T/ p.Val114Leu	de novo	AD	Novel	Campomelic dysplasia
58	Male	0.92	MCA	CREBBP	c.3306T>A/ p.Tyr1102*	de novo	AD	Novel	Rubinstein-Taybi syndrome

DD, developmental delay; ID, intellectual disability; MCA, multiple congenital anomalies; *Variants included in public databases or previously published literature.

TABLE 3 | Patients with pathogenic and likely pathogenic variants identified by copy number variation sequencing.

Case no.	Sex	Age (y)	Categories	CNV	CNV length	Diagnosed disorders
1	Male	0.17	MCA	seq[GRCh37]del(22)(22q11.2)chr22:g.18963217_21541547del	2.58Mb	22q11.2 deletion syndrome
3	Female	1	MCA	seq[GRCh37]del(5)(5p15.33p15.1)chr5:g.54215_16752126del	16.70Mb	Cri-du-chat syndrome
11	Female	0.58	DD/ID+MCA	seq[GRCh37]del(9)(9p24.3p11.2)chr9:g.152081_43829455del	43.68Mb	9p deletion syndrome
13	Male	8	ASD	seq[GRCh37]del(22)(22q11.21)chr22:g.19028192_21466206del	2.44Mb	22q11.2 deletion syndrome
20	Male	4	DD/ID+MCA	seq[GRCh37]del(11)(11q24.2q25)chr11:g.124175265_134974797del	10.80Mb	NA
23	female	0.17	DD/ID+MCA	seq[GRCh37]dup(8)(8q24.3)chr8:g.141319205_144918800dup; seq[GRCh37]del(18)(18q22.2q23)chr18:g.67971593_77882168del	3.60Mb; 9.91Mb	NA
28	Female	8	DD/ID+MCA	seq[GRCh37]del(7)(7q11.23)chr7:g.72645343_74161920del	1.52Mb	Williams-Beuren syndrome
29	Female	15	MCA	seq[GRCh37]del(14)(14q13.1q21.1)chr14:g.35268524_38367321del	3.10Mb	NA
36	Male	3	DD/ID+MCA	seq[GRCh37]del(4)(4q25q26)chr4:g.111697561_116176319del	4.48Mb	NA
59	Female	0.67	MCA	seq[GRCh37]del(22)(22q11.2)chr22:g.18905953_20318038del	1.41Mb	22q11.2 deletion syndrome

ASD, autism spectrum disorders; DD, developmental delay; ID, intellectual disability; MCA, multiple congenital anomalies.

disease and laryngomalacia. After diagnosis, these patients took further evaluation and treatment following the clinical practice guidelines for individuals with 22q11.2 deletion syndrome (Bassett et al., 2011).

DISCUSSION

In this study, we performed parallel tests of WES and CNV-seq for 60 patients. Thirty-seven pathogenic or likely pathogenic variants in 32 patients were identified, and the overall diagnosis rate was 53.3%. Our study provided preliminary results of clinical utility of parallel tests of CNVs and SNVs at the same time in pediatric patients with developmental delay/intellectual disability, autism spectrum disorders, and multiple congenital anomalies.

Over 6500 phenotypes were included in the OMIM by the end of 2019 and most of them were onset during childhood. Genetic pediatric disorders are highly heterogeneous and relatively rare, which, for investigating them, necessitates a process of serial testing for specific conditions (Han and Lee, 2020). In consequence, this strategy may be expensive and time-consuming. The molecular diagnosis of these rare diseases is based on the appropriate choice among tests of karyotype, Sanger sequencing, multiplex ligation dependent probe amplification, CMA, NGS, etc., which requires a physician with medical genetics training and experience. However, most children's hospitals do not have independent genetic clinic for patients with genetic disorders (Hu et al., 2018). Meanwhile, a more comprehensive strategy of genomic-first approach should be applied to achieve higher positive rate and reduce diagnostic odyssey (Johnson, 2015).

Previous WES studies revealed a diagnosis rate of ~25–30% in nonselective patients (Yang et al., 2013, 2014; Lee et al., 2014; Daoud et al., 2016; Retterer et al., 2016). Moreover, additional pathogenic and likely pathogenic CNVs were also identified in over 10% of patients (Sanmann et al., 2015; Homma et al., 2018; Jang et al., 2019). The development of NGS data based CNV calling algorithm makes it possible to identify SNVs and CNVs at same time. In this study, we performed trio-WES and low-coverage WGS (0.3×) simultaneously and uncovered pathological causes in over half of the patients. The rate of patients with neurological features was 67% (14/21), which was higher than previous reports (Gao et al., 2019; Jiao et al., 2019). However, it was notable that our sample size was relatively small, and more researches were needed to better the clinical efficacy.

The application of NGS to detect genome-wide CNVs is a recently developed method (Xie and Tammi, 2009). Comparing to pediatric disorders, CNV-seq was more widely used in prenatal screening and diagnosis, which required lower resolution (Liang et al., 2014; Zhang R. et al., 2019; Zhao and Fu, 2019). The resolution of CNV-seq can be adjusted by increasing or decreasing the sequencing data volume of raw data. For pediatric diagnosis, the resolution was usually >100 Kb (Gao et al., 2019; Jiao et al., 2019). By comparing differences of aligned reads number between case and control samples, losses or gains of chromosomal regions can be identified and quantitated.

Therefore, it has been reported that it is easier to identify low-level mosaicism by CNV-seq, rather than CMA (Grotta et al., 2015). Furthermore, identifying CNVs by sequencing has other advantages, including lower starting input of the DNA, cheaper cost, shorter TAT and higher throughput.

In the present study, our parallel tests of WES and low-pass WGS ($0.3\times$)-based CNV-seq decreased TAT and human labor without a significantly increased cost. Compared to WES, the extra cost was used for WGS of each sample to yield one Gb raw data directly, using the same sample library constructed during WES before capture, and the sequencing cost for 1 Gb raw data was about $10 USD with Illumina NovaSeq, in this research. We identified pathogenic or likely pathogenic variants with WES and CNV-seq in 36.7 and 16.7% of patients, respectively, which was comparable with previous study, indicating that our strategy is efficient to simultaneously identify SNVs and CNVs. Therefore, the ordering physicians do not need to choose between CMA and NGS, a choice which is dependent on clinical expertise, and they are more likely to identify clinically relevant variants in patients with atypical presentation or novel phenotypes. The mean age of positive cases at diagnosis was 3 years (95% CI: 1.73, 4.26). Early diagnosis avoided long diagnostic odyssey once symptoms appear, but also provided opportunities for early intervention, helping to improve outcomes and prolonging life.

In conclusion, our study provided preliminary experience of parallel tests of WES and CNV-seq with same initial library and evaluated the clinical utility. Compared to traditional trio-WES test, our strategy increased the diagnosis rate from 36.7 to 53.3% in our relatively small cohort. With less depending on physicians' selection between SNVs and CNVs tests, we propose this strategy is efficient and cost-effective for the diagnosis of genetic pediatric disorders with high heterogeneity.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available due to concerns regarding participant/patient anonymity. Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

Written informed consent was obtained from the individuals and their legal guardians for the publication.

AUTHOR CONTRIBUTIONS

XH, RG, JG, and ZQ collected the clinical data. CH and WL designed the study. XH, RG, and JG performed data analysis. XH and CH wrote the manuscript. All authors contributed to data acquisition and data interpretation, critically revised the manuscript for important intellectual content, approved the final, submitted version of the manuscript, and agreed to be accountable for all aspects of the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00473/full#supplementary-material

REFERENCES

Bassett, A. S., Mcdonald-Mcginn, D. M., Devriendt, K., Digilio, M. C., Goldenberg, P., Habel, A., et al. (2011). Practical guidelines for managing patients with 22q11.2 deletion syndrome. *J. Pediatr.* 159, 332-339 e331. doi: 10.1016/j.jpeds.2011.02.039

Bowling, K. M., Thompson, M. L., Amaral, M. D., Finnila, C. R., Hiatt, S. M., Engel, K. L., et al. (2017). Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med.* 9:43. doi: 10.1186/s13073-017-0433-1

Daoud, H., Luco, S. M., Li, R., Bareke, E., Beaulieu, C., Jarinova, O., et al. (2016). Next-generation sequencing for diagnosis of rare diseases in the neonatal intensive care unit. *CMAJ* 188, E254–E260. doi: 10.1503/cmaj.150823

Das, P., Stockton, D. W., Bauer, C., Shaffer, L. G., D'souza, R. N., Wright, T., et al. (2002). Haploinsufficiency of PAX9 is associated with autosomal dominant hypodontia. *Hum. Genet.* 110, 371–376. doi: 10.1007/s00439-002-0699-1

Gao, C., Wang, X., Mei, S., Li, D., Duan, J., Zhang, P., et al. (2019). Diagnostic Yields of Trio-WES accompanied by CNVseq for rare neurodevelopmental disorders. *Front. Genet.* 10:485. doi: 10.3389/fgene.2019.00485

Grotta, S., D'elia, G., Scavelli, R., Genovese, S., Surace, C., Sirleto, P., et al. (2015). Advantages of a next generation sequencing targeted approach for the molecular diagnosis of retinoblastoma. *BMC Cancer* 15, 841. doi: 10.1186/s12885-015-1854-0

Han, J. Y., and Lee, I. G. (2020). Genetic tests by next generation sequencing in children with developmental delay and/or intellectual disability. *Clin. Exp. Pediatr.* doi: 10.3345/kjp.2019.00808. [Epub ahead of print].

Hayashi, S., Yagi, M., Morisaki, I., and Inazawa, J. (2015). Identical deletion at 14q13.3 including PAX9 and NKX2-1 in siblings from mosaicism of unaffected parent. *J. Hum. Genet.* 60, 203–206. doi: 10.1038/jhg.2014.123

Homma, T. K., Krepischi, A. C. V., Furuya, T. K., Honjo, R. S., Malaquias, A. C., Bertola, D. R., et al. (2018). Recurrent copy number variants associated with syndromic short stature of unknown cause. *Horm. Res. Paediatr.* 89, 13–21. doi: 10.1159/000481777

Hu, X., Gui, B., Su, J., Li, H., Li, N., Yu, T., et al. (2017). Novel pathogenic ACAN variants in non-syndromic short stature patients. *Clin. Chim. Acta* 469, 126–129. doi: 10.1016/j.cca.2017.04.004

Hu, X., Li, N., Xu, Y., Li, G., Yu, T., Yao, R. E., et al. (2018). Proband-only medical exome sequencing as a cost-effective first-tier genetic diagnostic test for patients without prior molecular tests and clinical diagnosis in a developing country: the China experience. *Genet. Med.* 20, 1045–1053. doi: 10.1038/gim.2017.195

Hu, X., Liu, J., Guo, R., Guo, J., Zhao, Z., Li, W., et al. (2019). A novel 14q13.1–21.1 deletion identified by CNV-Seq in a patient with brain-lung-thyroid syndrome, tooth agenesis and immunodeficiency. *Mol. Cytogenet.* 12:51. doi: 10.1186/s13039-019-0463-z

Jang, W., Kim, Y., Han, E., Park, J., Chae, H., Kwon, A., et al. (2019). Chromosomal microarray analysis as a first-tier clinical diagnostic test in patients with developmental delay/intellectual disability, autism spectrum disorders, and multiple congenital anomalies: a prospective multicenter Study in Korea. *Ann. Lab. Med.* 39, 299–310. doi: 10.3343/alm.2019.39.3.299

Jiao, Q., Sun, H., Zhang, H., Wang, R., Li, S., Sun, D., et al. (2019). The combination of whole-exome sequencing and copy number variation sequencing enables the diagnosis of rare neurological disorders. *Clin. Genet.* 96, 140–150. doi: 10.1111/cge.13548

Johnson, N. E. (2015). Whole-exome sequencing in neurologic practice: reducing the diagnostic odyssey. *Neurol. Genet.* 1:e37. doi: 10.1212/NXG.0000000000000037

Kearney, H. M., Thorland, E. C., Brown, K. K., Quintero-Rivera, F., South, S. T., and Working Group of the American College of Medical Genetics Laboratory Quality Assurance Committee. (2011). American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet. Med.* 13, 680–685. doi: 10.1097/GIM.0b013e3182217a3a

Lee, H., Deignan, J. L., Dorrani, N., Strom, S. P., Kantarci, S., Quintero-Rivera, F., et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 312, 1880–1887. doi: 10.1001/jama.2014.14604

Liang, D., Peng, Y., Lv, W., Deng, L., Zhang, Y., Li, H., et al. (2014). Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J. Mol. Diagn.* 16, 519–526. doi: 10.1016/j.jmoldx.2014.05.002

Manning, M., Hudgins, L., Professional, P., and Guidelines, C. (2010). Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. *Genet. Med.* 12, 742–745. doi: 10.1097/GIM.0b013e3181f8baad

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., et al. (2010). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* 86, 749–764. doi: 10.1016/j.ajhg.2010.04.006

Monroe, G. R., Frederix, G. W., Savelberg, S. M., De Vries, T. I., Duran, K. J., Van Der Smagt, J. J., et al. (2016). Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet. Med.* 18, 949–956. doi: 10.1038/gim.2015.200

Nguyen, M. T., and Charlebois, K. (2015). The clinical utility of whole-exome sequencing in the context of rare diseases - the changing tides of medical practice. *Clin. Genet.* 88, 313–319. doi: 10.1111/cge.12546

Rajagopalan, R., Murrell, J. R., Luo, M., and Conlin, L. K. (2020). A highly sensitive and specific workflow for detecting rare copy-number variants from exome sequencing data. *Genome Med.* 12:14. doi: 10.1186/s13073-020-0712-0

Retterer, K., Juusola, J., Cho, M. T., Vitazka, P., Millan, F., Gibellini, F., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* 18, 696–704. doi: 10.1038/gim.2015.148

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). 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.* 17, 405–424. doi: 10.1038/gim.2015.30

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., et al. (2019). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG)

and the Clinical Genome Resource (ClinGen). *Genet. Med.* 22, 245–257. doi: 10.1038/s41436-019-0686-8

Rosati, A., Berti, B., Melani, F., Cellini, E., Procopio, E., and Guerrini, R. (2015). Recurrent drop attacks in early childhood as presenting symptom of benign hereditary chorea caused by TITF1 gene mutations. *Dev Med Child Neurol* 57, 777–779. doi: 10.1111/dmcn.12644

Sanmann, J. N., Pickering, D. L., Golden, D. M., Stevens, J. M., Hempel, T. E., Althof, P. A., et al. (2015). Assessing the utility of confirmatory studies following identification of large-scale genomic imbalances by microarray. *Genet. Med.* 17, 875–879. doi: 10.1038/gim.2014.204

Santen, G. W., Sun, Y., Gijsbers, A. C., Carre, A., Holvoet, M., Haeringen, A., et al. (2012). Further delineation of the phenotype of chromosome 14q13 deletions: (positional) involvement of FOXG1 appears the main determinant of phenotype severity, with no evidence for a holoprosencephaly locus. *J. Med. Genet.* 49, 366–372. doi: 10.1136/jmedgenet-2011-100721

Setter, S. M., Neumiller, J. J., Dobbins, E. K., Wood, L., Clark, J., Duvall, C. A., et al. (2009). Treatment of chorea associated with Huntington's disease: focus on tetrabenazine. *Consult. Pharm.* 24, 524–537. doi: 10.4140/tcp.n.2009.524

Stark, Z., Tan, T. Y., Chong, B., Brett, G. R., Yap, P., Walsh, M., et al. (2016). A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet. Med.* 18, 1090–1096. doi: 10.1038/gim.2016.1

Sun, Y., Ye, X., Fan, Y., Wang, L., Luo, X., Liu, H., et al. (2020). High detection rate of copy number variations using capture sequencing data: a retrospective study. *Clin. Chem.* 66, 455–462. doi: 10.1093/clinchem/hvz033

Vissers, L., Van Nimwegen, K. J. M., Schieving, J. H., Kamsteeg, E. J., Kleefstra, T., Yntema, H. G., et al. (2017). A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genet. Med.* 19, 1055–1063. doi: 10.1038/gim.2017.1

Wang, Z., Andrews, P., Kendall, J., Ma, B., Hakker, I., Rodgers, L., et al. (2016). SMASH, a fragmentation and sequencing method for genomic copy number analysis. *Genome Res.* 26, 844–851. doi: 10.1101/gr.201491.115

Xie, C., and Tammi, M. T. (2009). CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* 10:80. doi: 10.1186/1471-2105-10-80

Yang, Y., Muzny, D. M., Reid, J. G., Bainbridge, M. N., Willis, A., Ward, P. A., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* 369, 1502–1511. doi: 10.1056/NEJMoa1306555

Yang, Y., Muzny, D. M., Xia, F., Niu, Z., Person, R., Ding, Y., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 312, 1870–1879. doi: 10.1001/jama.2014.14601

Yao, R., Zhang, C., Yu, T., Li, N., Hu, X., Wang, X., et al. (2017). Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. *Mol. Cytogenet.* 10:30. doi: 10.1186/s13039-017-0333-5

Zhang, P., Hu, X., Guo, R., Guo, J., Li, W., Qian, S., et al. (2019). Novel HMGCS2 pathogenic variants in a Chinese family with mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase deficiency. *Pediatr. Investig.* 3, 86–90. doi: 10.1002/ped4.12130

Zhang, R., Chen, X., Wang, D., Chen, X., Wang, C., Zhang, Y., et al. (2019). Prevalence of chromosomal abnormalities identified by copy number variation sequencing in high-risk pregnancies, spontaneous abortions, and suspected genetic disorders. *J. Int. Med. Res.* 47, 1169–1178. doi: 10.1177/0300060518818020

Zhao, X., and Fu, L. (2019). Efficacy of copy-number variation sequencing technology in prenatal diagnosis. *J. Perinat. Med.* 47, 651–655. doi: 10.1515/jpm-2019-0005