Whole exome sequencing for developmental delay and learning difficulties: abridged secondary publication

SLJ Kwok, WLE Hau, TF Chan, FMI Lo, LWE Fung, SKW Tsui *

KEY MESSAGES

1. Whole exome sequencing (WES) was performed for 30 patient-parent trios of children with undiagnosed developmental delay.
2. Our bioinformatics pipeline coupled with clinical review confirmed the diagnosis by *de novo* mutations or inherited compound heterozygous rare variants in seven (24%) patients. In addition, possible causative variants were found in three other patients.
3. WES is a cost-efficient method for management of patients with apparently undiagnosed developmental delay or learning difficulties after first-tier testing.

Hong Kong Med J 2020;26(Suppl 7):S4-6
HMRF project number: 01120596

[1] SLJ Kwok, [2] WLE Hau, [3,4] TF Chan, [2] FMI Lo, [5] LWE Fung, [1,4] SKW Tsui

[1] *School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong*
[2] *Clinical Genetic Service, Department of Health, Hong Kong*
[3] *School of Life Sciences, The Chinese University of Hong Kong, Hong Kong*
[4] *Hong Kong Bioinformatics Centre, The Chinese University of Hong Kong, Hong Kong*
[5] *Department of Paediatrics, The Chinese University of Hong Kong, Hong Kong*

* Principal applicant and corresponding author: kwtsui@cuhk.edu.hk

Introduction

Developmental disorder (DD) encompasses developmental delay and learning difficulties. It is a disorder in which the most prominent pathogenetic mechanism occurs during embryogenesis or early brain development.[1] In Hong Kong, the total number of persons with intellectual disability was estimated to be 71 000 to 101 000, representing a prevalence of 1.0% to 1.4%. First-tier tests such as G-banding karyotyping and array chromosome genomic hybridisation have low diagnostic yield. Recently, whole-exome sequencing (WES) has been successful in diagnosing 30% of previously undiagnosed DDs.[2] We aim to find the causative, small, single-nucleotide or multi-nucleotide mutations fitting monogenic dominant disease models (such as *de novo* mutations [DNMs]) and recessive disease models.

Methods

Patients with undiagnosed DDs and their parents were recruited through Prince of Wales Hospital, The Chinese University of Hong Kong, and Clinical Genetic Service of the Department of Health in Hong Kong. They had negative results of first-tier tests (G-banding karyotyping and array chromosome genomic hybridisation). Those with potentially acquired causes to explain their DDs were excluded. Informed written consent was obtained from both parents for all patients who were under 18 years old. Clinical data of initial apparent phenotypes were collected. Peripheral blood was collected from each patient-parent trio. QIAamp DNA Blood Mini kit (Qiagen) was used to extract genomic DNA (gDNA) according to the manufacturer's protocol. The purified, quality-passed samples were sent to Beijing Genomics Institute for sequencing.

WES data were preprocessed according to the GATK Best Practices version 3. For the variant prioritisation, GEMINI was chosen as the SQL-based interface to first query the variants according to various inheritance patterns including *de novo* 'inheritance'. GEMINI requires Vt to preprocess the files. Vt was also used to decompose and normalise variants. The resulting normalised VCF file underwent functional annotation using the Ensembl Variant Effect Predictor version 84, with the option enabled for PolyPhen and SIFT predictions of deleteriousness of missense variants.

The annotated VCF file was then loaded into a SQL database using the GEMINI 'load' command. To find DNMs, the GEMINI '*de novo*' command was used. The number of DNMs per human exome is usually <5 and therefore no filtering criteria (eg, functional deleteriousness prediction) were applied. Besides, we were interested in any type of DNM in or flanking the exons. To find the variants fitting the compound heterozygous or autosomal homozygous recessive inheritance models, the GEMINI commands 'comp_hets' and 'autosomal_recessive' were used. For X-linked hemizygous variants, the GEMINI 'query' command was used, which allows one to construct flexible genotype queries that can be set to filter for said variants. For the final step of variant prioritisation via bioinformatics pipeline, we wrote a custom script in Perl to automatically

indicate whether a gene in the GEMINI query output matched a known DD gene according to the Developmental Disorders Genotype-to-Phenotype database, as well as the reported allelic requirement for the DD to manifest (ie, dominant/monoallelic or recessive/biallelic). The shortlisted variants per patient were then handed over to the clinicians of our team, who then reviewed previous case reports of DD patients with the same gene with clinically and/or experimentally confirmed causative mutation.

Results

A total of 30 patients with undiagnosed DDs and their parents were recruited. After basic quality filtering by VQSR and depth, the bottom approximately 30% of worst-quality or low-depth variants were removed. The Table shows patients with clinically confirmed and validated DNMs or compound heterozygous variants.

Patient 4 with SCN8A DNM was deemed causative. Her clinician prescribed sodium channel blocker (Trileptal) to improve her seizures but her seizures got worse again a few months later. The clinician also performed vagal nerve stimulation in December 2015, and the patient's condition appeared to stabilise.

Patient 28 with another SCN8A DNM had a relatively stable condition. The clinician did not change prescription after diagnosis, because the family requested so.

Patient 10 had CACNA1A DNM, which the clinician attributed to his chronic ataxia and recurrent hemiparesis. At least six mutations in this gene were reported to be associated with familial hemiplegic migraine and cerebellar signs, including T666M, R583Q, D715E, Y1385C, R1668W, and W1684R.[3] These can be explained by high expression of α1 subunit in the cerebellum, involving the Purkinje and granules cells.

Patient 14 was referred for delayed development with DNM in GRIN2B. There are studies supporting the contribution of GRIN2B to developmental issues and autistic features in children. Specifically, the stereotypical hand movement, which was a form of behavioural phenotype noticed in our patient, was delineated with the GRIN2B mutation.

In patient 16, the mutated GNAO1 gene is abundant in brain tissue and is believed to be important in brain function. Although patient 16 did not present with epilepsy, the clinical features of DD and involuntary movement could be explained with GNAO1 DNM.

Patient 20 has an X-linked ATRX variant inherited from his mother. His blood was taken for further non-genetic investigation. His clinical features were compatible with the diagnosis of X-linked alpha thalassemia mental retardation.

Patient 23 had multiple problems (in addition to learning difficulties) throughout years of follow-up in the Clinical Genetic Service. She had a progressive loss of vision, obesity, short stature, increased serum lipid level with fatty liver, and mild thyroid dysfunction. All these features were compatible with the diagnosis of Alstrom syndrome, which was causally related to ALMS1 mutations. The ALMS1 compound heterozygous frameshift

TABLE. Patients with clinically confirmed mutations

Patient of trio	Gene	Mutation inheritance type	mRNA mutation in HGVS notation	Protein mutation in HGVS notation	Mutation consequence type	Deleterious prediction by programs (damaging or neutral)				
						CADD	SIFT	Pol-Phen-2	PRO-VEAN	FATHMM
4	SCN8A	DNM	NM_014191.3:c.4862T>G	NP_055006.1:p.Leu1621Trp	Missense	Damaging	Damaging	Damaging	Damaging	Damaging
7	POLR1C	Compound heterozygous	NM_203290.3:c.436T>C	NP_976035.1:p.Cys146Arg	Missense	Damaging	Damaging	Damaging	Damaging	Damaging
	POLR1C	Compound heterozygous	NM_203290.3: c.883_885delAAG	NP_976035.1:p.Lys295del	In-frame deletion	Damaging	N/A	N/A	Damaging	N/A
10	CACNA1A	DNM	NM_000834.3:c.2065G>A	NP_000825.2:p.Gly689Ser	Missense	Damaging	Damaging	Damaging	Damaging	Damaging
14	GRIN2B	DNM	NM_000834.3:c.2065G>A	NP_000825.2:p.Gly689Ser	Missense	Damaging	Damaging	Damaging	Damaging	Neutral
16	GNAO1	DNM	NM_020988.2:c.736G>A	NP_066268.1:p.Gly246Lys	Missense	Damaging	Damaging	Damaging	Damaging	Damaging
20	ATRX	X-linked	ENST00000373344.5: c.536A>G	ENSP00000362441.4: p.Asn179Ser	Missense	Damaging	Damaging	Damaging	Damaging	Damaging
23	ALMS1	Compound heterozygous	ENST00000264448.6: c.4911_4914delTAAA	ENSP00000264448.6: p.Asn1637LysfsTer4	Frameshift	Damaging	N/A	N/A	N/A	N/A
	ALMS1	Compound heterozygous	ENST00000264448.6: c.11110_11128delAGGTCT AATCAAATTAAAA	ENSP00000264448.6: p.Arg3704LeufsTer11	Frameshift	Damaging	N/A	N/A	N/A	N/A
24	PPP2R5D	DNM	NM_006245.3:c.592G>A	NP_006236.1:p.Glu198Lys	Missense	Damaging	Damaging	Damaging	Damaging	Neutral
26	MAGEL2	DNM	NM_019066.4:c.187dupC	NP_061939.3:p. Gln63ProfsTer47	Frameshift	Damaging	N/A	N/A	N/A	N/A
28	SCN8A	DNM	NM_014191.3:c.434A>G	Np_055006.1:p.Asn145Ser	Missense	Damaging	Damaging	Damaging	Damaging	Damaging

mutations identified could fully support a valid genotype-phenotype correlation.

Patient 13 had congenital heart disease and poor muscle tone. The DNM found in TRRAP (transformation/transcription domain-associated protein), which is the cofactor of the histone acetyltransferase, might play a role in controlling cell-cycle progression and neurogenesis.

Patient 15 had moderate learning difficulties, extra digit, hydrocephalus, and dysmorphic facial features. These features might be related to the PRPF8 DNM. Individuals with this mutation may have various problems because PRPF8 gene variant was found to link to the regulatory potential of altering the core spliceosome machinery.

Patient 26 was clinically suspected to have Prader-Willi syndrome because of severe learning difficulties, marked hypotonia, small head, and relevant facial characteristics. Although MAGEL2 DNM is related to Prader-Willi syndrome, the truncating mutation in our patient differed from it.

Discussion

WES of patient-parent trios enables mutation calling according to inheritance patterns. Parents can act as healthy controls (at least clinically healthy although there is a possibility of subclinical symptoms of DD) and provide the best genetic background controls.

After basic quality filtering by VQSR and depth, the bottom 30% of worst-quality or low-depth variants were removed. This helped prioritise variants. For example, a DNM is a type of Mendelian violation, and therefore a considerable portion of the worst-quality variants are Mendelian violations. Basic variant filtering helps reduce the time spent on manually excluding false positives. Also, a lot of low-quality genotype variants are false positives and filtered out. False positives should be minimised during the bioinformatics analysis when possible. However, most filtering recommendations are conservative in order to avoid excluding true positive variants that might be of clinical significance.

In this study, the pipeline we developed is sufficient to make clinical diagnosis, with a success rate similar to previous small- and medium-scale exome studies. Recently, the ExAC consortium pointed out that many exome studies had been over-conservative in prioritising variants, especially on over-emphasising rare variants, leaving many potentially DD causative variants to be filtered out.[4] In another study of protein-coding genes that were almost depleted of loss-of-function variants, 70% of the genes were not associated with DDs.[5]

In our earlier pipeline, if the clinician deemed the DNM as not causative, then we looked for recessive variants. Later, we modified the pipeline so that the preliminarily prioritised variants of all inheritance types (ie, compound heterozygous, autosomal recessive, X-linked hemizygous) were queried together and collectively sent to the clinicians. Preliminary prioritisation by the research staff was appropriate because the clinicians only need to confirm variants with previous case reports.

For the clinical management, patient 4 was given new drug and palliative treatment according to his SCN8A DNM diagnosis. The discovery of CACNA1A DNM in patient 10 corrected the diagnosis of hemiplegic migraine. The diagnosis of patient 20 helped the clinician decide on follow-up after confirming already available clinical features. Other diagnosed patients mostly used the WES results for genetic counselling.

Conclusion

Making a genetic diagnosis for DD can avoid unnecessary investigations and enable the use of potentially more useful drugs (eg, sodium channel blocker), early consideration of palliative surgery (eg, vagal nerve stimulation), counselling on the risk of sudden unexpected death in epilepsy, and appropriate genetic counselling.

Funding

This study was supported by the Health and Medical Research Fund, Food and Health Bureau, Hong Kong SAR Government (#01120596). The full report is available from the Health and Medical Research Fund website (https://rfs1.fhb.gov.hk/index.html).

Disclosure

Part of the results of this research have been previously published in:

1. Fung LW, Kwok SL, Tsui KW. SCN8A mutations in Chinese children with early onset epilepsy and intellectual disability. Epilepsia 2015;56:1319-20.

References

1. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. Am J Hum Genet 2009;84:524-33.
2. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature 2015;519:223-8.
3. Ducros A, Denier C, Joutel A, et al. The clinical spectrum of familial hemiplegic migraine associated with mutations in a neuronal calcium channel. N Eng J Med 2001;345:17-24.
4. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.
5. Walsh R, Thomson KL, Ware JS, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. Genet Med 2017;19:192-203.