



OPEN ACCESS

# Resolving unsolved whole-genome sequencing data in paediatric neurological disorders: a cohort study

Ching-Shiang Chi,<sup>1</sup> Chi-Ren Tsai,<sup>1</sup> Hsiu-Fen Lee <sup>1,2</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/archdischild-2024-326985>).

<sup>1</sup>Division of Pediatric Neurology, Children's Medical Center, Taichung Veterans General Hospital, Taichung, Taiwan  
<sup>2</sup>Department of Post-Baccalaureate Medicine, College of Medicine, National Chung Hsing University, Taichung, Taiwan

## Correspondence to

Dr Hsiu-Fen Lee, Division of Pediatric Neurology, Children's Medical Center, Taichung Veterans General Hospital, Taichung, Taiwan; [lee hf@hotmail.com.tw](mailto:lee hf@hotmail.com.tw)

C-SC and C-RT contributed equally.

C-SC and C-RT are joint first authors.

Received 10 February 2024  
Accepted 11 May 2024

## ABSTRACT

**Objective** To resolve unsolved whole-genome sequencing (WGS) data in individuals with paediatric neurological disorders.

**Design** A cohort study method using updated bioinformatic tools, new analysis targets, clinical information and literature databases was employed to reanalyse existing unsolved genome data.

**Participants** From January 2016 to September 2023, a total of 615 individuals who aged under 18 years old, exhibited neurological disorders and received singleton WGS were recruited. 364 cases were unsolved during initial WGS analysis, in which 102 consented to reanalyse existing singleton WGS data.

**Results** Median duration for reanalysis after initial negative WGS results was 2 years and 4 months. The diagnostic yield was 29 of 102 individuals (28.4%) through reanalysis. New disease gene discovery and new target acquisitions contributed to 13 of 29 solved cases (44.8%). The reasons of non-detected causative variants during initial WGS analysis were variant reclassification in 9 individuals (31%), analytical issue in 9 (31%), new emerging disease–gene association in 8 (27.6%) and clinical update in 3 (10.3%). The 29 new diagnoses increased the cumulative diagnostic yield of clinical WGS in the entire study cohort to 45.5% after reanalysis.

**Conclusions** Unsolved paediatric WGS individuals with neurological disorders could obtain molecular diagnoses through reanalysis within a timeframe of 2–2.5 years. New disease gene, structural variations and deep intronic splice variants make a significant contribution to diagnostic yield. This approach can provide precise genetic counselling to positive reanalysis results and end a diagnostic odyssey.

## INTRODUCTION

Paediatric neurological disorders are grouped as diseases characterised by high levels of phenotypic heterogeneity and genotypic diversity. Conventional stepwise diagnostic strategies to proceed genetic tests are often protracted, costly and inconclusive.<sup>1</sup> With molecular evolution, next-generation sequencing (NGS), including targeted gene panels, whole-exome sequencing (WES) and whole-genome sequencing (WGS), has become the mainstream for uncovering genetic causes. WGS has the highest diagnostic yield and is more powerful than the other NGS approaches because it provides a comprehensive testing platform to uncover protein-coding variants, structural variations (SVs), non-coding variants, DNA repetition disorders and mitochondrial mutations.<sup>2 3</sup>

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Whole-genome sequencing (WGS) has been clinically demonstrated as a first-tier diagnostic tool for paediatric neurodevelopmental disorders.
- ⇒ Although WGS is more powerful than the other next-generation sequencing approaches, a multitude of patients remain undiagnosed.
- ⇒ Knowledge growth in genomic medicine has prompted the reanalysis of existing clinical genome data as a standard method in making further diagnoses.

## WHAT THIS STUDY ADDS

- ⇒ Diagnostic yield of clinical WGS data reanalysis in this study cohort outperforms other studies due to a longer analytical interval between our analyses.
- ⇒ New disease gene, structural variations and deep intronic splice variants make a significant contribution to diagnostic yield.
- ⇒ The cumulative diagnostic yield of clinical WGS in paediatric neurological disorders increases to 45.5% after reanalysis.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ A feasible timeframe for initial negative WGS data reanalysis to be 2–2.5 years dependent on clinical updates and bioinformatic advances is recommended.

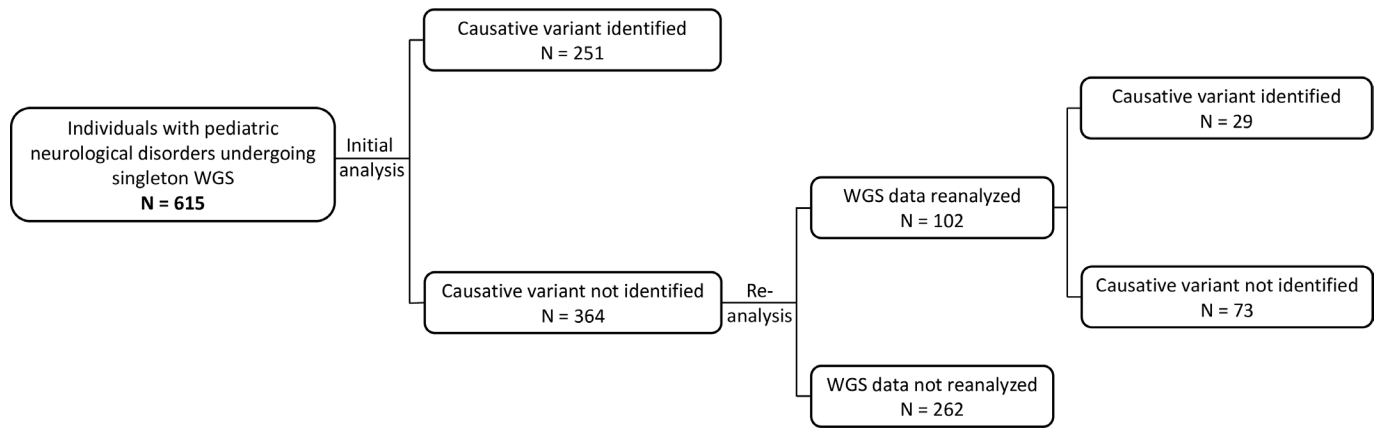
WGS was previously considered limited for research purposes due to high cost and innumerable variants with problematic interpretation.<sup>4</sup> Nowadays, with cost reduction and incomparable advantages, WGS has been clinically demonstrated as a first-tier diagnostic framework for paediatric neurology patients.<sup>3 5–10</sup> The diagnostic yield of clinical WGS varies widely, yet could achieve rates of 24.7–54%.<sup>6 7 9 11–15</sup> However, a multitude of patients remain undiagnosed. Evolutionary bioinformatics and knowledge growth in genomic medicine have prompted the reanalysis of existing clinical genome data as a standard method in making further diagnoses.<sup>13 16</sup>

Our article published in 2021 noted interpretational limitations for classification of certain genomic variants as pathogenic/likely pathogenic.<sup>14</sup> These may be existing insufficient evidence of American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Chi C-S, Tsai C-R, Lee H-F. *Arch Dis Child* Epub ahead of print: [please include Day Month Year]. doi:10.1136/archdischild-2024-326985



**Figure 1** Flow chart illustrating for individual recruitment process in reanalysis of unsolved singleton whole-genome sequencing (WGS) data.

AMP) 2015 guideline,<sup>17</sup> identification challenges for SVs, technical restrictions for unreachable coding sequence regions or uncertain functions of variants. The present extended cohort study incorporates 102 previously unsolved WGS cases. The diagnostic yield of clinical genome sequencing (GS) data reanalysis was 28.4%. New disease gene, SVs and non-coding region variants account for 44.8% of reanalysed positive individuals.

## METHODS

### Participant recruitment

This was a prospective successive investigation for our previous cohort study.<sup>14</sup> From January 2016 to September 2023, individuals younger than 18 years of age who had neurodevelopmental disorders were recruited for singleton WGS. Patients with definite molecular diagnoses by chromosomal microarray analysis (CMA) and/or single gene testing were excluded from this study. The details of inclusion/exclusion criteria for individuals undergoing singleton WGS have been described elsewhere.<sup>14</sup>

The cohort included 615 affected individuals (figure 1). In initial analysis, 251 individuals obtained molecular diagnoses from WGS; 364 probands, whose DNA had been analysed, and for whom a molecular diagnosis was not achieved at the time of the initial analysis, were categorised as unsolved. Unsolved cases were invited for the reanalysis of existing genomic data. Enrollment followed serial order, without specific inclusion/exclusion criteria, avoiding choice bias. 102 unsolved WGS individuals with parental consent were enrolled.

### Flow chart for reanalysis of unsolved clinical WGS

Sequence analysis using updated bioinformatic tools occurred in three phases: primary, secondary and tertiary (figure 2).

Primary analysis employed an Illumina sequencer.

Secondary analysis used FastaQ-formatted sequencing reads from the Illumina NovaSeq 6000 platform aligned to the human genome assembly (hg38) with DRAGEN V.4.0 to produce variant call format files for single nucleotide variant (SNV), small insertion and deletion (indel) and mitochondrial DNA (mDNA), and binary alignment map files for SVs. Secondary analysis of SVs was performed using a variety of software tools, including Manta (DRAGEN V.4.0), CNVnator v0.4.1, Delly v1.1.5 and SURVIVOR 1.0.7.

Tertiary analysis for variants of SNVs/indels and mDNA was annotated using the functional and effect prediction tools ANNOVAR 2020-06-08 and SnpEff v5.1d. Intronic variants of WGS data were predicted and filtered using SpliceAI with a >0.5 cut-off score.<sup>18</sup> Any variants showing a minor allele frequency

(MAF)  $\geq 1\%$  in any subpopulation database were excluded. Filtered variants were then considered for phenotypic association, or newly observable clinical characteristics, listed in the Online Mendelian Inheritance in Man (OMIM) database. Direct Sanger sequencing to identify specific DNA mutations confirmed potential findings. Similarly, Sanger sequencing, in combination with reverse transcription-PCR analysis, aided mRNA splicing variation confirmation. Additionally, allele segregation studies assisted confirmation of pathogenicity of the potential variations. The AnnotSV v3.2.3 performed SV tertiary analysis for further SV annotation. The potential disease-associated SVs were confirmed using array comparative genomic hybridisation, high-density single nucleotide polymorphism (HD SNP) array or long PCR, whereafter the segregation study was conducted.

Potential variation pathogenicity was classified using the InterVar v2.0.2, which is a tool that assesses pathogenicity of genetic variants according to the ACMG/AMP 2015 guidelines.<sup>17</sup> Positive molecular diagnosis was defined as relevant pathogenic/likely pathogenic variants correlated with clinical phenotypes and inheritance patterns, verification by Sanger sequencing or other molecular tests in the affected probands and their parents. A formal updated clinical report was issued.

### Data analysis

Diagnostic yield of clinical WGS data reanalysis and the reasons for undetected causative variants in the initial analysis were described.

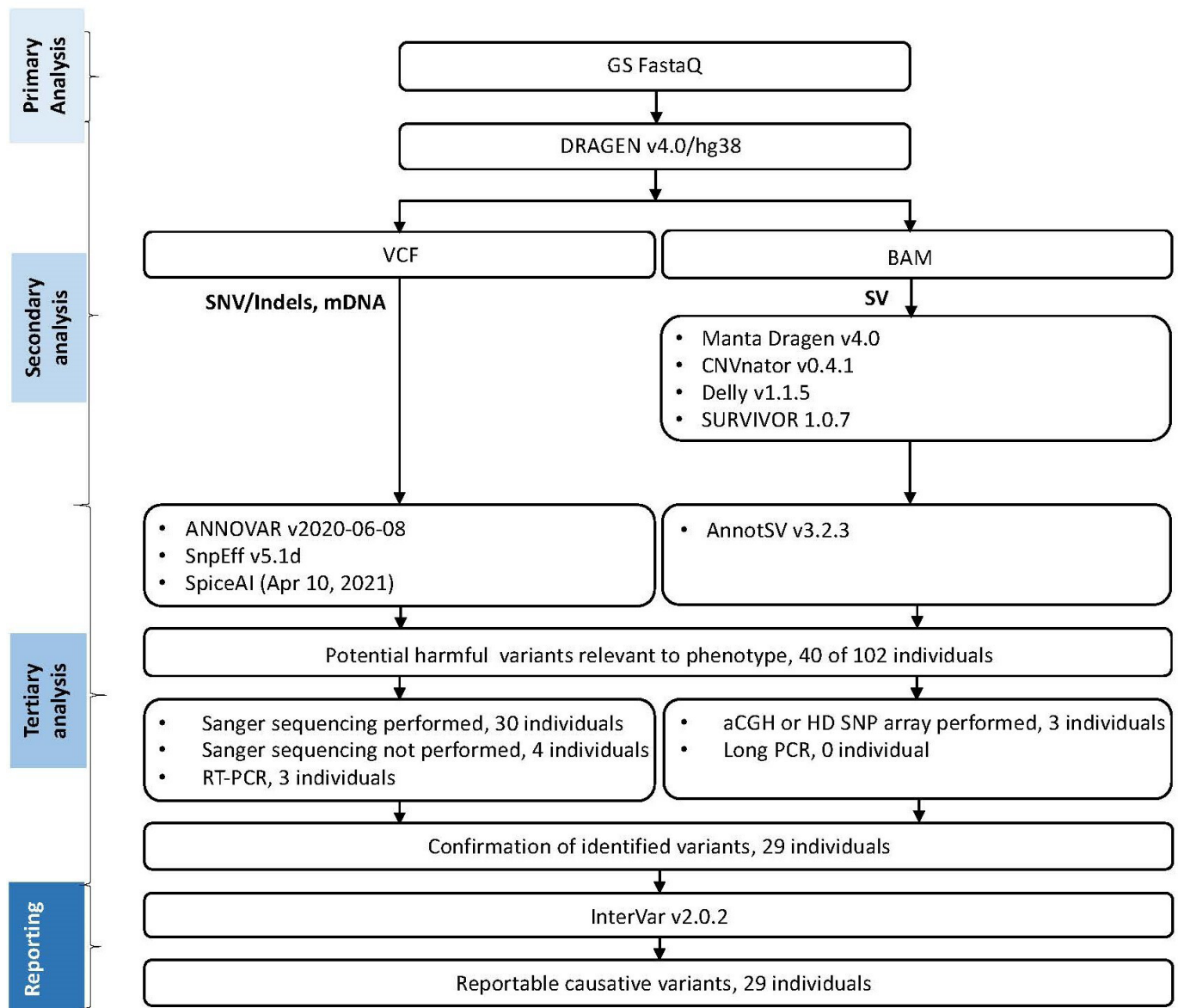
## RESULTS

### Diagnostic yield of clinical WGS data reanalysis

Median duration for reanalysis after initial negative WGS results was 2 years and 4 months, ranging from 2 months to 5 years. 29 unsolved WGS individuals received successful molecular diagnoses after reanalysis. The diagnostic rate of clinical WGS reanalysis was 28.4% (29/102). The 29 new diagnoses increased the cumulative diagnostic yield of clinical WGS in the entire study cohort from 40.8% (251/615) in the first analysis to 45.5% (280/615) after reanalysis. Table 1 summarises clinical phenotypes, identified gene variants, inheritance patterns and OMIM-based diagnostic names.

### Reasons for initial WGS variant non-detection

Online supplemental table S1 displays the four categories sorted for the reasons of non-detected causative variants during initial analysis: variant reclassification in 9 (9/29; 31%), analytical issue



**Figure 2** Schematic representation of manual reanalysis of singleton whole-genome sequencing data. Stepwise workflow for the primary, secondary and tertiary analyses, and reporting using updated versions of bioinformatic tools. aCGH, array-based comparative genomic hybridisation; BAM, binary alignment map; GS, genome sequencing; HD SNP, high-density single nucleotide polymorphism; indel, insertion/deletion; mDNA, mitochondrial DNA; RT-PCR, reverse transcription-PCR; SNV, single nucleotide variant; SV, structural variation; VCF, variant call format.

in 9 (9/29; 31%), new emerging disease–gene association in 8 (8/29; 27.6%) and clinical update in 3 (3/29; 10.3%).

#### Variant reclassification

Genome reanalysis reclassified genetic variants from uncertain significance (VUS) to pathogenicity/likely pathogenicity in nine individuals after incorporating clinical information, literature databases, bioinformatic tool updates and/or genetic characteristics.

Based on ACMG/AMP 2015 variant pathogenicity classification guidelines, ID 069 and ID 155 carried homozygous *TBC1D24* p.Ala500Val with VUS (PM5, PM2, PM1), with literature and ClinVar database overwhelming supporting this variant as pathogenic/likely pathogenic (PP5) and inherited in trans (PM3), thus reclassified accordingly as pathogenic. ID 122 had *SLC12A1* p.Lys908Glu with VUS (PP3, PP2, and PM2) but Bartter syndrome type 1 (PP4) indicative phenotype, where

genetic analysis uncovered this variant detected in trans (PM3), consequently reclassified as likely pathogenic. ID 174 and ID 175 had *SATB2* p.Glu402Lys with VUS (PS2 and PM2), with literature and ClinVar database supporting this variant robustly as pathogenic (PP5), thus reclassified as likely pathogenic. ID 207 had VUS of *GRIN2B* and ID 272 had VUS of *STXBP1* (PP3 and PM2), while revealing de novo trait after parental genetic analysis without family history (PS2), thus reclassified as likely pathogenic. ID 214's *TBCD* p.Pro1122Leu with VUS (PP3 and PM2) was reclassified as likely pathogenic after incorporating evidence literature and ClinVar database supporting this variant as pathogenic (PP5) and genetic analysis (PM3). ID 323 with VUS of *HCN1* (PM2) employed new lines of computational findings substantiating a deleterious effect (PP3) with genetic analysis revealing this variant's de novo nature (PS2) and reclassified as likely pathogenic.

**Table 1** Phenotypes, identified gene names, inheritance patterns, causative variants and reasons for initial variant non-detection of the 29 new diagnoses after reanalysis

New diagnosis after reanalysis	N=29*
Phenotypes, n (%)	
Developmental delay/intellectual disability	25 (86.2)
Epilepsy	11 (37.9)
Dystonia	8 (27.6)
Neuropathy	3 (10.3)
Failure to thrive/short stature	2 (6.9)
Myopathy	1 (3.4)
Renal tubular acidosis	1 (3.4)
Spastic paraplegia	1 (3.4)
Stroke	1 (3.4)
Visual impairment	1 (3.4)
Gene name (n)	
Autosomal recessive disorder	<i>SHQ1</i> (8), <i>TBC1D24</i> (2), <i>GLS</i> (1), <i>SLC12A1</i> (1), <i>SLC25A3</i> (1), <i>TBCD</i> (1), <i>TTN</i> (1)
Autosomal dominant disorder	<i>MNF2</i> (2), <i>SATB2</i> (2), <i>BPTF</i> (1), <i>CREBBP</i> (1), <i>FGF12</i> (1), <i>GRIN2B</i> (1), <i>HCN1</i> (1), <i>MORC2</i> (1), <i>RNF213</i> (1), <i>STXBP1</i> (1), <i>TRIP12</i> (1), <i>TUBB4</i> (1)
Mode of inheritance	
Autosomal recessive disorder, n (%)	15 (51.7)
Compound heterozygosity	12
Homozygous mutations	3
Autosomal dominant disorder, n (%)	14 (48.3)
De novo	12
Inherited	2
Causative variants	
Single nucleotide variant, n (%)	26 (89.7)
Exonic variant	24
Deep intronic splicing variant	2
Structural variation, n (%)	3 (10.3)
Causative variant alleles	
N=44	
Genetic novelty, n (%)	29 (65.9)
Autosomal recessive disorder	24
Autosomal dominant disorder	5
Mutation type, n (%)	
Missense	28 (63.6)
Nonsense	7 (15.9)
Duplication	4 (9.1)
Deletion	3 (6.8)
Splicing	2 (4.5)
Reasons for variants not detected in initial WGS	
N=29	
Variant reclassification, n (%)	9 (31)
Analytical issue, n (%)	9 (31)
New emerging disease–gene association, n (%)	8 (27.6)
Clinical update, n (%)	3 (10.3)

\*Additional information appears in online supplemental table S1. WGS, whole-genome sequencing.

### Analytical issue

Nine individuals had analytical limitations in the initial analysis.

Initially, six patients lacked detectable causative variants. Among them, three individuals were found to harbour SVs on reanalysis, confirmed using HD SNP array: ID 051 had 24% mosaicism for a 4.25 Mb duplication on chromosome

16p13.3, involving the *CREBBP* gene, which is associated with Rubinstein-Taybi syndrome; ID 291 showed an 859 kb microdeletion of chromosome 2q36.3q37.1, encompassing the *TRIP12* gene, which causes intellectual disability with/without autism spectrum disorders, speech delay and dysmorphic features; and ID 374 revealed a 594 kb duplication of chromosome 3q28q29 resulting in the majority duplication of the *FGF12* gene, which causes developmental and epileptic encephalopathy 47.

The other three individuals carried causative SNVs after reanalysis: ID 171's *TUBB4* p.Glu410Lys was classified as pathogenic through pathogenicity reanalysis (PP3 and PM2), literature database (PP5 and PS3) and genetic analysis (PS2); ID 283's *MORC2* p.Ser87Leu was classified as pathogenic after pathogenicity reanalysis (PP3 and PM2), literature database (PM1, PP5 and PS3) and genetic analysis (PS2); ID 335's *BPTF* p.Arg841Ter was similarly classified as pathogenic after pathogenicity reanalysis (PVS1 and PM2) and genetic analysis (PS2).

ID 498 initial genetic analysis identified a heterozygous nonsense mutation *TTN* p.Glu10383Ter (PS1). When new symptoms developed, further genetic analysis revealed an additional heterozygous variant in *TTN* p.Tyr27744Cys, which had a deleterious effect (PP3), was found to be rare in general population (PM2), and was reclassified as likely pathogenic due to maternal inheritance (PP4 and PM3). The originally detected *TTN* p.Glu10383Ter variant was also determined pathogenic due to paternal inheritance (PP4 and PM3).

Two individuals carried a heterozygous SNV at coding region matched with recessive disorder phenotypes during initial GS analysis, and another pathogenic heterozygous splicing variant was identified in the deep intron region at GS reanalysis by SpliceAI. Initial GS of ID 342 discovered *GLS* p.Gly262Asp missense variant-related developmental and epileptic encephalopathy 71, and genome reanalysis identified another *GLS* c.736-406A>G deep intronic splicing variant (online supplemental figure S1) that inherited in trans; and that of ID 384 identified causative variant *SLC25A13* p.Ala554GlyfsTer17-related neonatal-onset citrullinemia type II while genome reanalysis found the deep intronic *SLC25A13* c.934-1926A>G splicing mutation that inherited in trans.

### New emerging disease–gene association

Genome reanalysis diagnosed eight individuals as biallelic *SHQ1* variant-related neurodevelopmental disorder with dystonia and seizures (OMIM 619922).

In January 2018, ID 007 carrying compound heterozygous *SHQ1* p.Leu333Val and *SHQ1* p.Tyr65Ter variants exhibited profound hypotonia and paroxysmal dystonia. ID 228, ID 419 and ID 434 carried the same genotype identified subsequently. Additionally, two siblings (ID 108 and ID 109) presented paroxysmal dystonia having compound heterozygous variants *SHQ1* p.Leu333Val and *SHQ1* p.Val271Glu. ID 335 possessed homozygous *SHQ1* p.Leu333Val variants and ID 451 had compound heterozygous *SHQ1* p.Leu333Val and p.Leu49Ser variants. The MAF for the four *SHQ1* variants in an East Asian population was <0.01% (PM2). Bioinformatic tools predicted the deleterious nature of these four *SHQ1* variants (PP3). Genetic analysis showed that the *SHQ1* gene variation followed autosomal recessive inheritance (PM3) and impaired normal *SHQ1* protein function as indicated through in vitro studies (PS3). In 2022, this phenotype was attributed to recessive *SHQ1* causative genes by OMIM.

### Clinical update

New diagnoses in three individuals occurred as a result of more apparent age-related, clinical phenotypes or family history updates.

ID 169, who showed sole neurological features (seizures and stroke) at the age of 1 year, with subsequent cardiac involvement (hypertrophic cardiomyopathy and hypertension), progressive cerebrovascular stenosis and bilateral renal arterial stenosis as she aged 6 years, was diagnosed with a de novo *RNF213* mutation-related moyamoya disease 2 (PP4, PM1, PM4, PM2 and PS2).

ID 234 was diagnosed with Charcot-Marie-Tooth disease, axonal, type 2A2A, caused by a heterozygous variant of *MFN2* after reanalysis. This variant was identified by genetic analysis in both ID 234 and his mother at initial WGS analysis, and it was classified as VUS (PP3, PM5 and PM2) because the mother had cryptic phenotypes. Reanalysis and detailed family history review indicated that the mother had, starting aged 20, experienced frequent falls. Neurological examination revealed high arched feet, clumsiness walking and deep tendon reflex absence. Maternal data update led to molecular diagnosis (PP4) and literature database (PP5). Through the same diagnostic process as ID 234, ID 299 inherited the corresponding *MFN2* variant from the phenotypically cryptic father.

### DISCUSSION

WGS is mainly used in patients with unsolved WES in clinical practice. Therefore, reanalysis of the original WGS data is rarely mentioned. The detection rates of diagnostic variants from WGS reanalysis in paediatric patients range from 4.2% (2/48)<sup>16</sup> to 10.9% (7/64).<sup>19</sup> In the present study, reanalysis of clinical WGS data yielded 28.4% (29/102) new diagnoses. Our clinical WGS reanalysis outperforms other studies, possibly due to the 6-year interval between our analyses, significantly longer than the analytical interval in other studies. This study exemplifies our reanalysis procedure's value in improving WGS diagnosis rates. A feasible timeframe for initial negative WGS data reanalysis to be 2–2.5 years dependent on clinical updates and bioinformatic advances is recommended.

SVs are believed to play a major role in the phenotype of different diseases, but such variation has been difficult to uniformly identify and characterise from the large number of human genomes because their identification is hindered by technical challenges intrinsic to short-read-based high-throughput sequencing technologies.<sup>20</sup> Submicroscopic chromosomal SVs account for about 15–20% of paediatric patients with neurodevelopmental disorders, where CMA was recognised as the first-line test.<sup>21</sup> However, WGS reportedly has higher diagnostic performance and clinical physicians tend to prioritise WGS over CMA in healthcare decision-making to find out causative variants more than SV.<sup>6 22</sup> The workflow proposed in this study combines multiple SV calling algorithms to mitigate obstacles to SV detection of single algorithm.

A WGS approach excels at uncovering pathogenic variants in non-coding regions.<sup>23</sup> SpliceAI is an open-source deep-learning algorithm and it has demonstrated a high ability to predict DNA variation-caused splicing defects.<sup>18 24</sup> The two individuals shown in this study carried deep intronic variants located in the -406 and -1926 coding regions of the two causative genes, which were far from the initial context of genome interpretation in potentially significant variants for the  $\pm 10$  splicing region. Therefore, it is imperative to consider intronic splicing variants in causative genes, especially in individuals with clinical phenotypes matching a specific recessive disorder, yet only

carry a single identified heterozygous variant, or in individuals with unsolved GS.<sup>25</sup>

Numerous rare and newly identified emerging disease-causing genes have been discovered during the genomic era. The variants in the same gene contributing to a disease trait may already have been identified in multiple, unrelated patients affected with similar phenotypes, and are then followed up with functional studies to provide evidence of gene causality. Therefore, a period of time could be needed to define a new disease according to the causative genes in OMIM, resulting in a time lag to the diagnosis, as shown the scenario in recessive *SHQ1* variant-related neurodevelopmental disorder.<sup>26–29</sup>

Regardless of the rationale for identifying causative variants through genome reanalysis, clinical updates cannot be overemphasised. Our study highlighted that when both probands and their parents carried the same gene variants in an autosomal dominant manner, clinical physicians should reassess the clinical and family histories to assist the clinical laboratory in the validation of variant pathogenicity. Moreover, initial genome-negative raw data reanalysis is needed when the phenotypes of individuals evolve.

Although the clinical WGS cumulative diagnostic yield for the entire study cohort was 45.5% after reanalysis, there remained limitations in this study. A significant number of GS-negative samples still escape identification after reanalysis due to unidentified causative GS genes, gene variations beyond current NGS technology scope or incomplete optimisation of variation analysis software. Sequencing long fragment DNA with third-generation sequencing may enable the analysis of gene variations in patients who are currently undiagnosable with short-read GS.

### CONCLUSIONS

WGS data reanalysis could diminish risks of false-negative reports, hindering the translation of genomic discoveries into clinical diagnosis and potentially increasing reproductive risks. Overcoming these challenges requires periodic analysis of unsolved WGS data by updating databases and accurate in-hospital communications. Automated genome analysis (re)sequencing is urgently needed to alleviate laboratory workload burden.

**Contributors** C-SC participated in recruitment of patients, acquisition, analysis and interpretation of data. C-RT participated in manual reanalysis and interpretation of singleton genome sequencing data. H-FL made great contributions to recruitment of patients and interpretation of data, revised the manuscript critically for important intellectual content and gave final approval of the version. H-FL is guarantor.

**Funding** This work was supported by Rare Disease Prevention and Treatment Subsidy Program, Health Promotion Administration, Ministry of Health and Welfare, Taiwan (grant number 2021).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Taichung Veterans General Hospital (TCVGH IRB CE20022A and TCVGH IRB CE22134B).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,

terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Hsiu-Fen Lee <http://orcid.org/0000-0002-6411-0183>

#### REFERENCES

- 1 Roberts JL, Hovanes K, Dasouki M, *et al*. Chromosomal microarray analysis of consecutive individuals with autism spectrum disorders or learning disability presenting for genetic services. *Gene* 2014;535:70–8.
- 2 Vissers LELM, van Nimwegen KJM, Schieving JH, *et al*. A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genet Med* 2017;19:1055–63.
- 3 Nisar H, Wajid B, Shahid S, *et al*. Whole-genome sequencing as a first-tier diagnostic framework for rare genetic diseases. *Exp Biol Med (Maywood)* 2021;246:2610–7.
- 4 Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. *Clin Med* 2017;17:545–51.
- 5 Bowdin SC, Hayeems RZ, Monfared N, *et al*. The Sickkids Genome Clinic: developing and evaluating a pediatric model for individualized genomic medicine. *Clin Genet* 2016;89:10–9.
- 6 Runheim H, Pettersson M, Hammarsjö A, *et al*. The cost-effectiveness of whole genome sequencing in neurodevelopmental disorders. *Sci Rep* 2023;13:6904.
- 7 Lionel AC, Costain G, Monfared N, *et al*. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med* 2018;20:435–43.
- 8 Marshall CR, Chowdhury S, Taft RJ, *et al*. Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med* 2020;5:47.
- 9 Lindstrand A, Eisfeldt J, Pettersson M, *et al*. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. *Genome Med* 2019;11:68.
- 10 Sun Y, Liu F, Fan C, *et al*. Characterizing sensitivity and coverage of clinical WGS as a diagnostic test for genetic disorders. *BMC Med Genomics* 2021;14:102.
- 11 Stranneheim H, Lagerstedt-Robinson K, Magnusson M, *et al*. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med* 2021;13:40.
- 12 Davis RL, Kumar KR, Puttick C, *et al*. Use of whole-genome sequencing for mitochondrial disease diagnosis. *Neurology* 2022;99:e730–42.
- 13 Bowling KM, Thompson ML, Amaral MD, *et al*. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med* 2017;9:43.
- 14 Lee HF, Chi CS, Tsai CR. Diagnostic yield and treatment impact of whole-genome sequencing in paediatric neurological disorders. *Dev Med Child Neurol* 2021;63:934–8.
- 15 Wright CF, Campbell P, Eberhardt RY, *et al*. Genomic diagnosis of rare pediatric disease in the United Kingdom and Ireland. *N Engl J Med* 2023;388:1559–71.
- 16 James KN, Clark MM, Camp B, *et al*. Partially automated whole-genome sequencing reanalysis of previously undiagnosed pediatric patients can efficiently yield new diagnoses. *NPJ Genom Med* 2020;5:33.
- 17 Richards S, Aziz N, Bale S, *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;17:405–24.
- 18 Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, *et al*. Predicting splicing from primary sequence with deep learning. *Cell* 2019;176:535–48.
- 19 Costain G, Jobling R, Walker S, *et al*. Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. *Eur J Hum Genet* 2018;26:740–4.
- 20 Chaisson MJP, Sanders AD, Zhao X, *et al*. Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat Commun* 2019;10:1784.
- 21 Türkylmaz A, Geckinli BB, Tekin E, *et al*. Array-based comparative genomic hybridization analysis in children with developmental delay/intellectual disability. *Balkan J Med Genet* 2021;24:15–24.
- 22 Clark MM, Stark Z, Farnaes L, *et al*. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* 2018;3:16.
- 23 Zhang F, Lupski JR. Non-coding genetic variants in human disease. *Hum Mol Genet* 2015;24:R102–10.
- 24 de Sainte Agathe J-M, Filser M, Isidor B, *et al*. SpliceAI-visual: a free online tool to improve SpliceAI splicing variant interpretation. *Hum Genomics* 2023;17:7.
- 25 Moore AR, Yu J, Pei Y, *et al*. Use of genome sequencing to hunt for cryptic second-hit variants: analysis of 31 cases recruited to the 100000 Genomes Project. *J Med Genet* 2023;60:1235–44.
- 26 Bizarro J, Meier UT. Inherited SHQ1 mutations impair interaction with Nap57/dyskerin, a major target in dyskeratosis congenita. *Mol Genet Genomic Med* 2017;5:805–8.
- 27 Sleiman S, Marshall AE, Dong X, *et al*. Compound heterozygous variants in SHQ1 are associated with a spectrum of neurological features, including early-onset dystonia. *Hum Mol Genet* 2022;31:614–24.
- 28 Chi CS, Tsai CR, Lee HF. Biallelic SHQ1 variants in early infantile hypotonia and paroxysmal dystonia as the leading manifestation. *Hum Genet* 2023;142:1029–41.
- 29 Rasmussen SA, Hamosh A, Amberger J, *et al*. What's in a name? Issues to consider when naming Mendelian disorders. *Genet Med* 2020;22:1573–5.