Genome-wide variant calling in reanalysis of exome sequencing data uncovered a pathogenic TUBB3 variant

Boer, E. de; Yaldiz, B.; Denomme-Pichon, A.S.; Matalonga, L.; Laurie, S.; Steyaert, W.; ... ; Solve-RD-DITF-ITHACA

European Journal of Medical Genetics 65 (2022) 104402

Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: www.elsevier.com/locate/ejmg

Genome-wide variant calling in reanalysis of exome sequencing data uncovered a pathogenic *TUBB3* variant

Elke de Boer [a], Burcu Yaldiz [b], Anne-Sophie Denommé-Pichon [c,d], Leslie Matalonga [e], Steve Laurie [e], Solve-RD SNV-indel working group, Wouter Steyaert [b], Rick de Reuver [b], Christian Gilissen [b], Michael Kwint [a], Rolph Pfundt [a], Solve-RD-DITF-ITHACA, Alain Verloes [f,g], Michèl A.A.P. Willemsen [h], Bert B.A. de Vries [a], A. Vitobello [c,d], Tjitske Kleefstra [a,1], Lisenka E.L. M. Vissers [a,*,1], Solve-RD SNV-indel working group, Solve-RD-DITF-ITHACA

[a] *Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Nijmegen, the Netherlands*
[b] *Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Nijmegen, the Netherlands*
[c] *Unité Fonctionnelle d'Innovation diagnostique des maladies rares, FHU-TRANSLAD, CHU Dijon Bourgogne, Dijon, France*
[d] *UFR Des Sciences de Santé, INSERM-Université de Bourgogne UMR1231 GAD «Génétique des Anomalies du Développement», FHU-TRANSLAD, Dijon, France*
[e] *CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Baldiri Reixac 4, Barcelona, 08028, Spain*
[f] *Département de Génétique, Hôpital Robert DEBRE, 48 boulevard Serurier, 75019, Paris, France*
[g] *INSERM UMR 1141 "NeuroDiderot", Hôpital R DEBRE, Paris, France*
[h] *Department of Pediatric Neurology, Amalia Children's Hospital, Radboud University Medical Centre, Donders Centre for Brain, Cognition, and Behaviour, Geert Grooteplein 10, 6525 GA, Nijmegen, the Netherlands*

ARTICLE INFO

Keywords:
TUBB3
Exome sequencing
Genome-wide variant calling
Solve-RD
ERN ITHACA

ABSTRACT

Almost half of all individuals affected by intellectual disability (ID) remain undiagnosed. In the Solve-RD project, exome sequencing (ES) datasets from unresolved individuals with (syndromic) ID (n = 1,472 probands) are systematically reanalyzed, starting from raw sequencing files, followed by genome-wide variant calling and new data interpretation. This strategy led to the identification of a disease-causing *de novo* missense variant in *TUBB3* in a girl with severe developmental delay, secondary microcephaly, brain imaging abnormalities, high hypermetropia, strabismus and short stature. Interestingly, the *TUBB3* variant could only be identified through reanalysis of ES data using a genome-wide variant calling approach, despite being located in protein coding sequence. More detailed analysis revealed that the position of the variant within exon 5 of *TUBB3* was not targeted by the enrichment kit, although consistent high-quality coverage was obtained at this position, resulting from nearby targets that provide off-target coverage. In the initial analysis, variant calling was restricted to the exon targets ± 200 bases, allowing the variant to escape detection by the variant calling algorithm. This phenomenon may potentially occur more often, as we determined that 36 established ID genes have robust off-target coverage in coding sequence. Moreover, within these regions, for 17 genes (likely) pathogenic variants have been identified before. Therefore, this clinical report highlights that, although compute-intensive, performing genome-wide variant calling instead of target-based calling may lead to the detection of diagnostically relevant variants that would otherwise remain unnoticed.

1. Introduction

Approximately 1% of the population is affected by intellectual disability (ID) (Maulik et al., 2011), but as the disorder is genetically and clinically highly heterogeneous, individual genetic causes are very rare, which complicates the diagnostic process and the identification of new disease-gene associations (Vissers et al., 2016). As a result, ~50–70% of individuals with ID remain undiagnosed (Srivastava et al., 2019). To overcome the

* Corresponding author.
E-mail address: Lisenka.Vissers@radboudumc.nl (L.E.L.M. Vissers).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ejmg.2021.104402
Received 1 September 2021; Accepted 28 November 2021
Available online 1 December 2021

problem of rarity and aiming to provide individuals with a diagnosis, the Solve-RD project (Zurek et al., 2021) comprises long-term sharing and systematic reanalysis of next generation sequencing and phenotypic data of 19,000 unsolved rare disease cases from four European Reference Networks (ERNs), including ~5,000 cases with (syndromic) ID from ERN-ITHACA (Intellectual Disability, TeleHealth and Congenital Anomalies; https://ern-ithaca.eu/). In contrast to recent studies on unsolved cohorts that reanalyzed variant calling files (Satterstrom et al., 2020) or pooled (*de novo*) variants (Kaplanis et al., 2020), in Solve-RD, new mapping and genome-wide variant calling is the starting point for reanalysis. Reanalysis strategies using genome-wide calling potentially increase diagnostic yield, as enrichment kits are often incomplete, despite the efforts aimed at the enrichment of all protein coding sequences (Ballester et al., 2016). By not restricting variant calling to (extended) targets of enrichment, the intrinsic imperfection of enrichment strategies to generate off-target coverage (Ballester et al., 2016) can be exploited to rescue some of the missing regions.

TUBB3 encodes the neuron-specific β-tubulin-3. Pathogenic *TUBB3* missense variants alter microtubule stability (Poirier et al., 2010; Tischfield et al., 2010) and impair axonal guidance in brain development (Huang et al., 2018), causing a spectrum of neuronal migration disorders (Poirier et al., 2010; Tischfield et al., 2010; Huang et al., 2018; Dentici et al., 2020; Fukumura et al., 2016; Shimojima et al., 2016; Whitman et al., 2016). Pathogenic variants can clinically and functionally be distinguished, and either increase microtubule stability, causing congenital fibrosis of extraocular muscles type 3A (MIM#600638) (Tischfield et al., 2010), or impair microtubule stability, leading to the highly variable phenotype of complex cortical dysplasia with other brain malformations type 1 (MIM#614039) (Poirier et al., 2010). A merge of both phenotypes has also been reported (Whitman et al., 2016).

We describe a disease-causing *TUBB3* variant identified in systematic reanalysis of exome sequencing (ES) data under the collaborative efforts of the Solve-RD project and ERN-ITHACA. Our clinical report illustrates the importance of starting reanalysis with raw sequencing data and stresses the diagnostic potential of genome-wide variant calling, especially in reanalysis.

2. Materials (subjects) and methods

2.1. Patient inclusion

In compliance with the local ethical guidelines and the Declaration of Helsinki, all individuals (or legal representatives) in the Solve-RD project provided informed consent. The Radboudumc institutional review board approved the study (2018–4986). For publication of photos, additional consent was obtained.

2.2. Initial diagnostic ES analysis

In 2014, diagnostic ES was performed locally as described previously (de Ligt et al., 2012). Of essence for this clinical report, exome enrichment was performed using the Agilent V4 Sure Select All Exons (GRCh37/Hg19) enrichment kit followed by sequencing on a HiSeq2000 Sequencing System (Illumina) to an anticipated median 75-fold sequence coverage on target regions. Variant calling was performed using the GATK HaplotypeCaller (v3.4) (Poplin et al., 2017), restricted to genomic coordinates of the exome targets±200bp.

2.3. Solve-RD data-sharing for systematic reanalysis

Human Phenotype Ontology (HPO)-standardized phenotypic data and BAM files of ES were uploaded to the RD-Connect Genome-Phenome Analysis Platform (GPAP; https://platform.rd-connect.eu/), and deposited at European Genome-Phenome Archive (EGA; accession numbers EGAZ00001527579, EGAZ00001534566 and EGAZ00001549096 for index, father and mother respectively), as part of the Solve-RD infrastructure (Zurek et al., 2021; Matalonga et al., 2021).

2.4. Variant identification

Systematic reanalysis of ES data in Solve-RD is described by Matalonga et al. (2021). The pipeline includes genome-wide variant calling (i.e. not restricted to the targets of the ES enrichment kit) with GATK (v3.6) (Poplin et al., 2017) to the GRCh37/Hg19 genome build, i.e. hs37d5, as used by Phase2 of the 1000 GP. Variants in genes associated with ID were annotated and prioritized by applying the following filtering parameters: 1) coverage depth ≥8 and genotype quality ≥20, 2) gnomAD (Karczewski et al., 2019) allele frequency <1%, 3) internal RD-Connect allele frequency <2%, 4) "Likely pathogenic" (ACMG, class 4) or "Pathogenic" (ACMG, class 5) in ClinVar (Landrum, 2018). Prioritized variants were interpreted in the context of the phenotype and visualized in the Integrative Genomics Viewer (IGV) (Robinson, 2011) alongside the exome data of the parents.

2.5. Sanger and amplicon-based deep sequencing

Variant validation by Sanger sequencing was performed using a routine diagnostic workflow (de Ligt et al., 2012). To exclude low-level mosaicism in the father, amplicon-based deep-sequencing was performed on the Ion Torrent platform from Life Technologies (Gilissen et al., 2014). PCR amplicons were generated in accordance with standard protocols, followed by read processing and alignment using JSI SeqNext Version 5.1.0 Build 503 (FW: GTGCGGAAGGAGTGTGAAAA; RV GGCATGAAGAAGTGCAGG).

2.6. Detecting off-target coverage on coding sequence and intersection with HGMD variants

Occurrence of off-target coverage on coding sequence across individuals was analyzed for enrichment kits Agilent V4 and V5 (Sure Select All Exons; GRCh37/Hg19), as ~70% of ES datasets shared for ERN ITHACA in the Solve-RD consortium was generated after enrichment with Agilent V4 or V5. We used 20 representative ES datasets for both kits, and defined robust off-target coverage occurring consistently across individuals at depth sufficiently for variant calling if 95% of tested individuals showed coverage exceeding 20-fold at coding sequence (Gencode v31 (Frankish et al., 2019)) located outside extended targets of the respective enrichment kit (genomic coordinates of the exome targets±200bp of Agilent V4 and V5). For the resulting regions, an intersection was made with disease-causing (DM) or likely disease-causing (DM?) SNVs/indels reported in HGMD Professional (Stenson et al., 2003) with specific attention for genes associated to ID (Radboudumc diagnostic ES ID gene panel DG3.1.0, comprising 1,392 genes) (Radboudumc. Intellectual, 2021).

3. Results

3.1. Clinical characteristics

We report on a 16-year-old female proband with severe ID, for whom extensive diagnostic testing in a university medical center in 2014 could not identify the (genetic) diagnosis. The proband was the first child from healthy, non-consanguineous parents and had a healthy brother and sister. Besides hyperemesis gravidarum, pregnancy was uneventful. Spontaneous vaginal delivery occurred at 42 weeks of gestation, and although delivery was complicated by meconium-stained amniotic fluid, there were no problems in the neonatal period (Apgar scores 9/9/10 after 1/5/10 min respectively; birth weight 3,480 g, −0.7 SD). At three months of age, developmental delay was evident, characterized by delayed motor and communicative milestones with notably quiet and undemanding behavior. Over the course of the first year, in which she was hospitalized for pneumonia and hypovolemic shock resulting from rotavirus gastro-enteritis, she gradually developed secondary microcephaly (head circumference at birth 0 SD, at 1 year −2 SD). Brain MRI (age 2 years and 2 months; Fig. 1) showed hypoplasia of the corpus callosum, a reduced volume of supratentorial white matter with delayed myelination, dysplasia of basal ganglia, thalami and cerebellum, and an enlarged ventricular system, mainly of the lateral

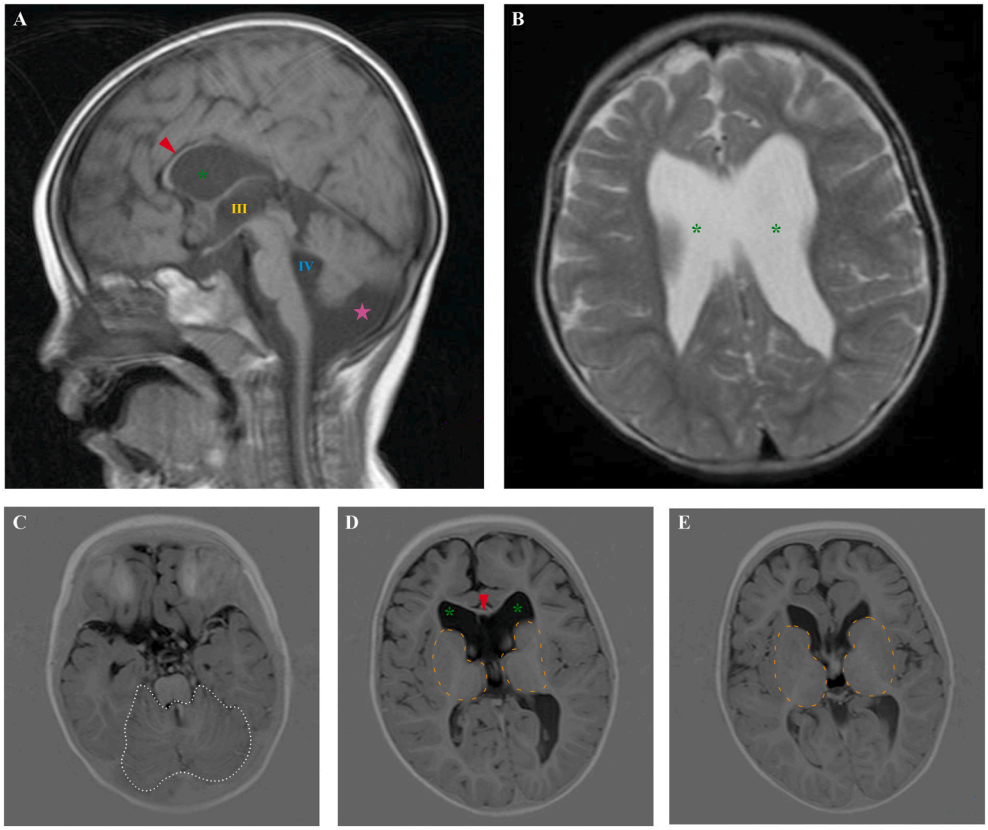

Fig. 1. Brain MRI displays structural brain abnormalities. Sagittal (A) and axial (C–E) T1-weighted and axial T2-weighted (B) MR images of the brain at age 2 years and 2 months show several structural brain abnormalities. There is general enlargement of the ventricular system, mostly affecting the lateral ventricles (indicated by green asterisks; panel A, B, D), but also the third ventricle (indicated by a yellow III; panel A) and the fourth ventricle (indicated by a blue IV; panel A), communicating with a mega-cisterna magna (indicated by a magenta star; panel A). The corpus callosum is hypoplastic (red arrowhead; panel A, D) and the cerebellum and cerebellar vermis show dysplasia, with abnormal morphology and orientation of folia (dotted white line; panel C). The basal ganglia and thalami are dysplastic, with an abnormal round appearance, and individual structures cannot be distinguished clearly (dashed orange line; panel D, E). Additionally, there is a reduced volume of supratentorial white matter (globally seen in panel B, D, E). . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

ventricles (left more prominently enlarged than right), but also of the third and fourth ventricle, with the fourth ventricle communicating with an enlarged cisterna magna. There were no signs of cortical malformation. Electromyography and nerve conduction studies (age 2 years) did not show signs of an associated peripheral neuropathy or myopathy. There was no epilepsy. At ages 11 and 15 years, she was tested by SON-R 2.5–7, a non-verbal intelligence test, which showed a severe developmental delay and intellectual disability (chronological age 11 years and 11 months: developmental age of 2 years and 10 months; chronological age 15 years and 9 months: developmental age 2 years and 11 months, developmental age in fluid reasoning 3 years, developmental age at performance scale 2 years and 9 months). At 15 years of age, she had a vocabulary of 10 words, with a more developed receptive vocabulary than expressive vocabulary (WPPSI-III-NL 4–7;11 subscale General Language Composite (GLC) at chronological age 15 years and 9 months: developmental age 5 years and below 2 years and 7 months respectively). Other medical problems included high hypermetropia (+7D/+7.5D), convergent strabismus, short stature (−2.5SD), scoliosis, susceptibility to urinary and upper respiratory tract infections, constipation, factor V Leiden thrombophilia and behavioral problems (automutilation, sleep disturbances, phonophobia). She had minor facial dysmorphisms, consisting of a round face, with deep-set eyes and large palpebral fissures, a low nasal bridge and a prominent chin (Fig. 2), a mild pectus carinatum and

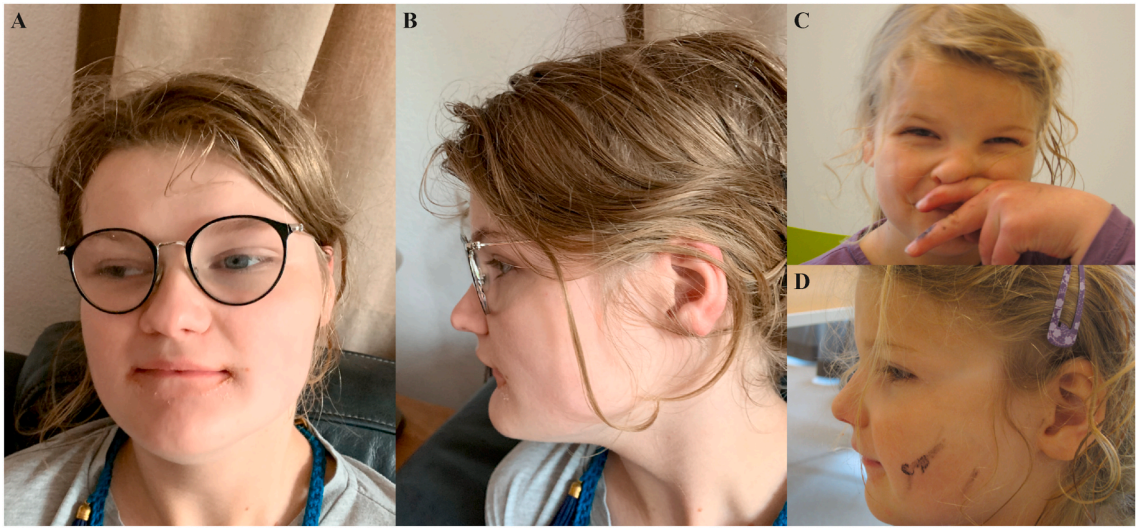

Fig. 2. Facial photographs of the proband. Frontal (panel A, C) and profile (panel B, D) facial photographs of the proband at age 17 years and 4 months (panel A, B) and at age 8 years and 2 months (panel C, D), showing mild facial dysmorphisms in the proband: she has a round face with deep-set eyes and large palpebral fissures, a low nasal bridge and a prominent chin.

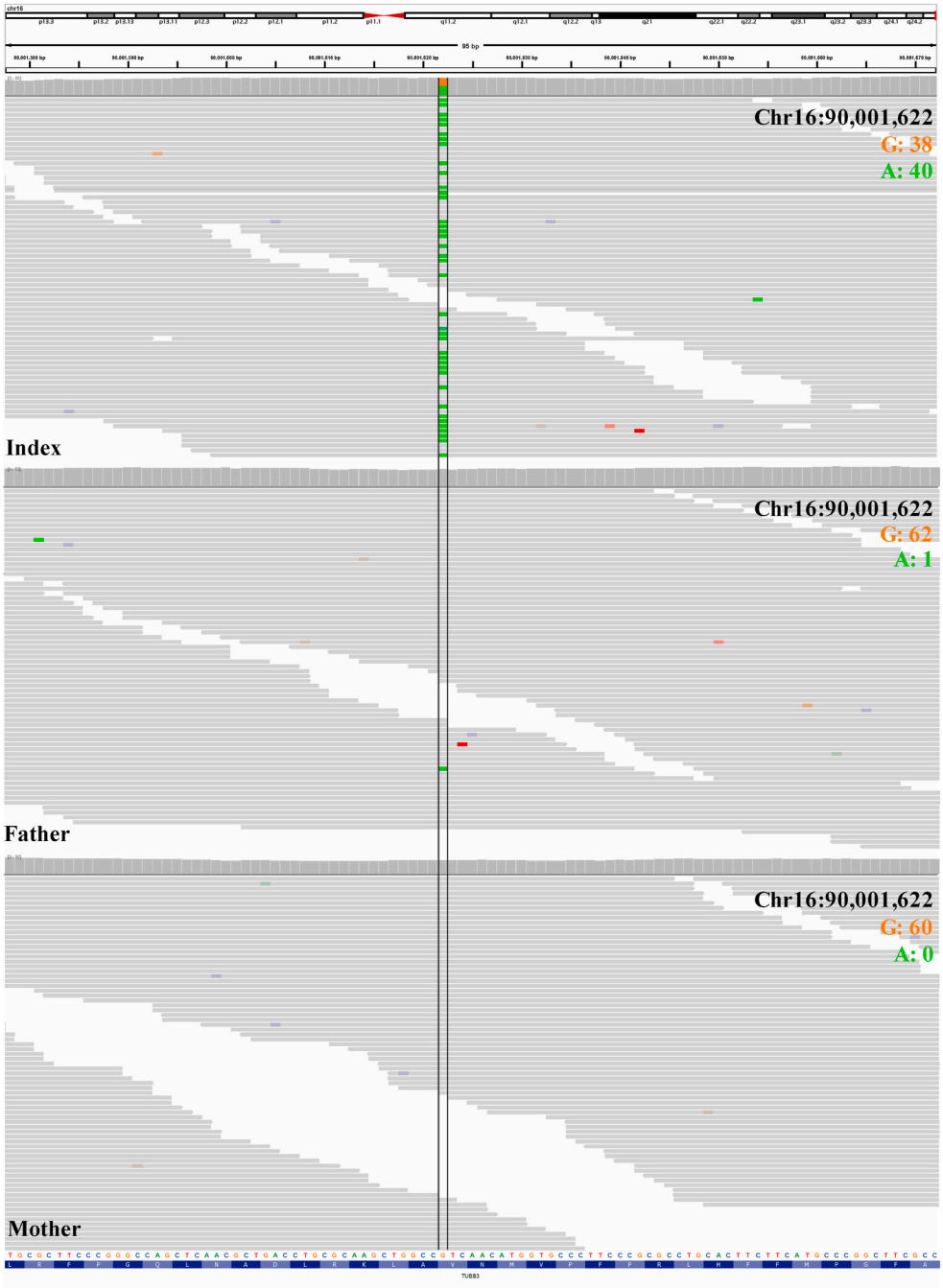

Fig. 3. IGV visualization of alignment files of index, father and mother. IGV visualization of alignment files of index (top), father (middle) and mother (bottom), showing the chr16:g.90001622G>A variant in the heterozygous state in the index (ref 38; alt 40), in one sequencing read of the father (ref 62; alt 1) and in no reads of the mother (ref 60; alt 0).

nipple retraction. Diagnostic evaluation consisted of karyotyping (at the age 1 of year), metabolic screening (age 1 year and 2 years), biochemical analysis of cerebrospinal fluid and blood (age 2 years), subtelomeric MLPA (age 2 years), 250K SNP array (age 4 years), targeted *GRIN2B* Sanger analysis (age 10 years), and trio-based ES (age 10 years), but did not result in an explanation for the phenotype.

3.2. Systematic reanalysis revealed a likely pathogenic TUBB3 variant

Data of the proband and parents were included in the Solve-RD project. Variant calling and annotation with the Solve-RD pipeline revealed a heterozygous *de novo* variant in *TUBB3*:Chr16(GRCh37/Hg19):g.90001622G>A;NM_006086.3:c.763G>A;p.(Val255Ile) in the proband. The *TUBB3* missense variant was absent from gnomAD (v2.1.1) (Karczewski et al., 2019), had a CADD-PHRED score of 24.8 (Rentzsch et al., 2019) and *TUBB3* is highly intolerant to variation (Z-score missense 4.85; pLI 0.97) (Karczewski et al., 2019), all supporting variant pathogenicity. In addition, the variant was previously reported as "Likely pathogenic" in ClinVar (VCV000372654.1)(Landrum, 2018), but without any clinical description. Reanalysis did not identify other clinically relevant *(de novo)* variants.

Manual inspection of sequence alignment files in IGV clearly showed the presence of the variant in the proband (Fig. 3), which was subsequently confirmed by Sanger sequencing. Sanger analysis of both parental samples indicated the variant to be of *de novo* origin. However, the BAM file of the father showed the same variant in 1 of 63 reads (Fig. 3),

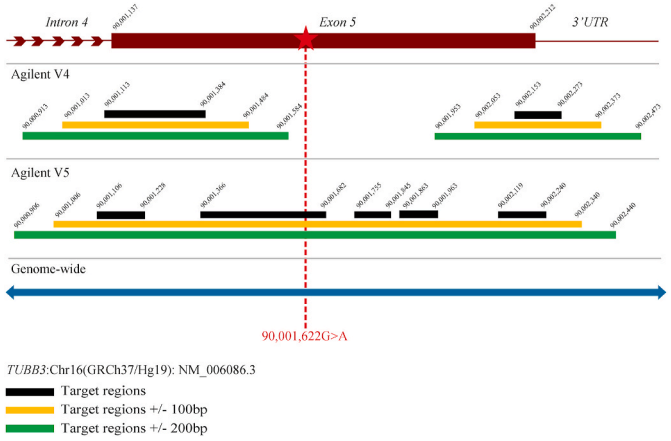

Fig. 4. Exon 5 of *TUBB3* and (extended) target regions used in variant calling. A schematic representation of the last part of *TUBB3,* including the variant identified in the proband. Shown in black are target regions of Agilent V4 and V5 enrichment kits, in orange target regions extended by 100bp on each side (the regions where sequence can be expected) and in green target regions extended by 200bp on each side (the regions used for variant calling in initial diagnostic analysis). The blue bar at the bottom visualizes genome-wide variant calling as done in the Solve-RD project. . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

potentially indicative of low-level mosaicism (Wright et al., 2019). Due to the technical limitations of Sanger sequencing to detect such low-level mosaicism (Acuna-Hidalgo et al., 2015), we therefore continued with amplicon-based deep sequencing of the *TUBB3* variant on DNA derived from blood of the father to a sequence depth of 11,000-fold. This confirmed the absence of the variant in father (G: 99.89%; A: 0.05%; T: <0.01%; C: 0%, with percentages of A, T and C equal to background noise resulting from PCR artefacts), and its *de novo* origin in the proband.

3.3. The variant calling algorithm in initial ES analysis did not detect the *TUBB3* variant

Next, we retrospectively analyzed why the *TUBB3* variant was not identified before, as *TUBB3* was included in the targets of the enrichment kit used for ES in this trio, and the gene was already described in literature in relation to intellectual disability and cortical dysplasia for 4 years at the time of initial diagnostic ES (Poirier et al., 2010). Additionally, in the diagnostic workflow used for individuals with intellectual disability, the interpretation of detected *de novo* variants in trio-based ES was performed invariably and with priority. Therefore, we first re-examined the output of the variant calling pipeline (annotated VCF) used in initial analysis and found that the variant was not called by the variant calling algorithm. Assessment of the regions considered in variant calling, showed that the variant is located in a region of the alignment not included in the extended targets for variant calling (target ± 200bp), and therefore ignored by the variant calling algorithm (Fig. 4). Hence, we concluded that the combination of target-based calling and the exome enrichment kit used explains why the initial diagnostic ES did not lead to the detection of the *TUBB3* variant.

3.4. Off-target covered regions disregarded in targeted variant calling coincide with known ID genes and previously reported pathogenic variants

To assess the overall impact of our conclusions, we checked the targets for *TUBB3* in Agilent V5, showing improved capture of *TUBB3,* with five additional targets compared to V4, of which three target exon 5 including the region of the variant (Fig. 4).

Next, we analyzed whether other regions of the coding sequence show off-target coverage but are disregarded in variant calling for both Agilent V4 and V5 enrichment kits (Table 1). First, we determined which protein coding sequences showed robust off-target coverage (≥20-fold), consistently present across multiple individuals (95% of 20 individuals), and then created the overlap with regions not included in targeted variant calling based on the extended targets (±200bp). This analysis detected a total of 840,347 and 761,594 nucleotides (1.95% and 1.76% of Gencode v31) in 1,537 and 1,298 of such regions, affecting 1,031 and 659 genes for Agilent V4 and V5, respectively. This included 36 established ID genes, including *TUBB3* for V4, and 21 ID genes for V5. By intersecting the results of V4 with HGMD Professional (Stenson et al., 2003), we found that 262 unique (likely) pathogenic SNVs/indels (DM? 33; DM: 229) affect 17 of these 36 ID genes. To a lesser extent, this was also observed for V5, where we showed 71 (likely) pathogenic variants (DM? 6; DM: 65) to be located in these regions, affecting 8 of 21 ID genes (Table 1, Supplementary Table 1).

4. Discussion

Through systematic reanalysis in the Solve-RD project, we found a likely pathogenic missense variant in *TUBB3* (c.763G>A;p.(Val255Ile)) in a 16-year-old female proband with severe ID. The variant is causative to the proband's phenotype, which consists of severe developmental delay, strabismus, microcephaly and structural brain abnormalities and fits the

Table 1
Genes and reported (likely) pathogenic variants in off-target covered regions disregarded in targeted variant calling.

Kit	Regions with ≥20x coverage in 95% of individuals	All genes			Genes associated with ID		
		# unique genes	# genes with previously reported (likely) pathogenic variants in these regions	# of reported unique (likely) pathogenic variants in these regions	# ID genes (% of all known ID genes)	# ID genes with previously reported (likely) pathogenic variants in these regions	# of reported unique (likely) pathogenic variants in these regions
Agilent V4	1,537	1,031	133	884	36 (2.6%)	17	262
Agilent V5	1,298	659	58	426	21 (1.5%)	8	71
Overlap V4 and V5		420	34	94	16 (1.1%)	5	13

phenotypic spectrum of complex cortical dysplasia with other brain malformations-1 (CDCBM1; MIM#614039) (Poirier et al., 2010; Dentici et al., 2020; Fukumura et al., 2016; Shimojima et al., 2016). Although the proband does not exhibit the typical cortical dysplasia phenotype, the peculiar combination of several of the other structural abnormalities seen in brain imaging of the proband, specifically dysplasia of the cerebellum, thalami and basal ganglia, hypoplasia of the corpus callosum, and enlargement of the ventricular system, is common in patients affected by pathogenic *TUBB3* missense variants (Dentici et al., 2020).

This variant, which explains the phenotype in this individual, could only be identified by subjecting sequencing data to genome-wide variant calling. Diagnostic variant calling pipelines were limited to detect this variant in exon 5 of *TUBB3*, as exon 5 was not completely targeted by the Agilent V4 enrichment kit, and the prior pipeline was dependent on targeted-based variant calling±200bp. Of note, the decision to use±200bp extended target calling was already double the recommendation of±100bp done by Agilent.

Despite not being fully targeted by the capture kit, exon 5 of *TUBB3* in this proband-parents trio showed sufficient off-target coverage, with the variant clearly detectable in the proband, enabling the genome-wide variant caller in the Solve-RD pipeline to identify the variant. To determine whether our observation was an isolated case, or whether this phenomenon may occur for multiple ID genes, we assessed the impact at exome-wide level. Exome-wide analysis for two enrichment kits across 20 individuals showed that interpretable off-target coverage in protein-coding sequence is generated for >1,000 genes in V4, and ~650 genes in V5. The use of target-based variant calling algorithms (±200bp) would disregard at least part of the coding sequence of these genes. The clinical relevance of these genes is highlighted by the fact that for multiple ID genes disease-causing variants similar to this variant in *TUBB3* have been reported in regions escaping variant calling. The impact may even be underestimated as we have used targets±200bp in diagnostic variant calling, already calling more off-target bases than recommended.

As reanalysis of existing ES data is a process used more often for individuals who remained undiagnosed at first diagnostic ES trajectory, the use of genome-wide variant calling should be considered. In particular if data were generated using 'older' exome enrichment kits, as it is to be expected that more recent enrichment kits keep improving and aim to resolve the issue reported here. Whether genome-wide variant calling should be used as a generalized approach for exome data during the first diagnostic analysis, is more difficult to assess as this depends on multiple factors, including the capture kit used and the (computational) infrastructure available.

In summary, we describe a female proband diagnosed with a *de novo* missense variant in *TUBB3,* through systematic, large-scale reanalysis on a cohort of unexplained (syndromic) ID cases within Solve-RD. We show that performing new variant calling enables detection of variants missed by variant calling in initial analysis, stressing the relevance of sharing raw data, as alternative approaches of data-sharing and reanalysis would have failed to identify these (potentially causative) variants. As the same issue also occurs for other (ID) genes, especially in older enrichment kits, this implies that genome-wide variant calling, although computationally demanding, expands the potential of ES and may effectuate new diagnoses in other unexplained cases of ERN ITHACA.

Funding

This work was financially supported by Aspasia grants of the Dutch Research Council (015.014.036 to TK and 015.014.066 to LELMV) and Netherlands Organization for Health Research and Development (917.183.10 to TK). The Solve-RD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 779257.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

We are very grateful to all families participating in the Solve-RD project and in particular to the proband and parents described in this clinical report. The collaborations in this study were facilitated by ERN ITHACA [EU Framework Partnership Agreement ID: 3HP-HP-FPA ERN-01-2016/739516, one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States, co-funded by the European Commission. For more information about the ERNs and the EU health strategy please visit https://ec.europa.eu/health/ern.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2021.104402.

References

Acuna-Hidalgo, R., Bo, T., Kwint, M.P., van de Vorst, M., Pinelli, M., Veltman, J.A., et al., 2015. Post-zygotic point mutations are an underrecognized source of de novo genomic variation. Am. J. Hum. Genet. 97 (1), 67–74.

Ballester, L.Y., Luthra, R., Kanagal-Shamanna, R., Singh, R.R., 2016. Advances in clinical next-generation sequencing: target enrichment and sequencing technologies. Expert Rev. Mol. Diagn 16 (3), 357–372.

de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., et al., 2012. Diagnostic exome sequencing in persons with severe intellectual disability. N. Engl. J. Med. 367 (20), 1921–1929.

Dentici, M.L., Maglione, V., Agolini, E., Catena, G., Capolino, R., Lanari, V., et al., 2020. TUBB3 E410K syndrome: case report and review of the clinical spectrum of TUBB3 mutations. Am. J. Med. Genet.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., et al., 2019. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. 47 (D1). D766-d73.

Fukumura, S., Kato, M., Kawamura, K., Tsuzuki, A., Tsutsumi, H., 2016. A mutation in the tubulin-encoding TUBB3 gene causes complex cortical malformations and unilateral hypohidrosis. Child Neurol Open 3, 2329048x16665758.

Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., et al., 2014. Genome sequencing identifies major causes of severe intellectual disability. Nature 511 (7509), 344–347.

Huang, H., Yang, T., Shao, Q., Majumder, T., Mell, K., Liu, G., 2018. Human TUBB3 mutations disrupt netrin attractive signaling. Neuroscience 374, 155–171.

Kaplanis, J., Samocha, K.E., Wiel, L., Zhang, Z., Arvai, K.J., Eberhardt, R.Y., et al., 2020. Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nature 586 (7831), 757–762.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., et al., 2019. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv, 531210.

Landrum, M.J., et al., 2018. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx1153.

Matalonga, L., Hernandez-Ferrer, C., Piscia, D., Schüle, R., Synofzik, M., Töpf, A., et al., 2021. Solving patients with rare diseases through programmatic reanalysis of genome-phenome data. Eur. J. Hum. Genet.

Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T., Saxena, S., 2011. Prevalence of intellectual disability: a meta-analysis of population-based studies. Res. Dev. Disabil. 32 (2), 419–436.

Poirier, K., Saillour, Y., Bahi-Buisson, N., Jaglin, X.H., Fallet-Bianco, C., Nabbout, R., et al., 2010. Mutations in the neuronal ß-tubulin subunit TUBB3 result in malformation of cortical development and neuronal migration defects. Hum. Mol. Genet. 19 (22), 4462–4473.

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., et al., 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv, 201178.

Radboudumc. Intellectual Disability Gene Panel DG 3.1.0 (1,392 Genes), 2021 [Available from. https://www.radboudumc.nl/getmedia/69d3205d-2e9b-49ed-bbc1-2b796fa146a9/INTELLECTUALDISABILITY_DG310.aspx.

Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., Kircher, M., 2019. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 47 (D1). D886-d94.

Robinson, et al., 2011. Integrative genomics viewer. Nat. Biotechnol. https://doi.org/10.1038/nbt.1754.

Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.Y., et al., 2020. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. Cell 180 (3), 568–584 e23.

Shimojima, K., Okamoto, N., Yamamoto, T., 2016. A novel TUBB3 mutation in a sporadic patient with asymmetric cortical dysplasia. Am. J. Med. Genet. 170a (4), 1076–1079.

Srivastava, S., Love-Nichols, J.A., Dies, K.A., Ledbetter, D.H., Martin, C.L., Chung, W.K., et al., 2019. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. Genet. Med. 21 (11), 2413–2421.

Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shiel, J.A., Thomas, N.S., et al., 2003. Human gene mutation database (HGMD): 2003 update. Hum. Mutat. 21 (6), 577–581.

Tischfield, M.A., Baris, H.N., Wu, C., Rudolph, G., Van Maldergem, L., He, W., et al., 2010. Human TUBB3 mutations perturb microtubule dynamics, kinesin interactions, and axon guidance. Cell 140 (1), 74–87.

Vissers, L.E., Gilissen, C., Veltman, J.A., 2016. Genetic studies in intellectual disability and related disorders. Nat. Rev. Genet. 17 (1), 9–18.

Whitman, M.C., Andrews, C., Chan, W.M., Tischfield, M.A., Stasheff, S.F., Brancati, F., et al., 2016. Two unique TUBB3 mutations cause both CFEOM3 and malformations of cortical development. Am. J. Med. Genet. 170a (2), 297–305.

Wright, C.F., Prigmore, E., Rajan, D., Handsaker, J., McRae, J., Kaplanis, J., et al., 2019. Clinically-relevant postzygotic mosaicism in parents and children with developmental disorders in trio exome sequencing data. Nat. Commun. 10 (1), 2985.

Zurek, B., Ellwanger, K., Vissers, L., Schüle, R., Synofzik, M., Töpf, A., et al., 2021. Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. Eur. J. Hum. Genet.