



## ARTICLE

# A Solve-RD ClinVar-based reanalysis of 1522 index cases from ERN-ITHACA reveals common pitfalls and misinterpretations in exome sequencing



### ARTICLE INFO

#### Article history:

Received 14 March 2022

Received in revised form

12 January 2023

Accepted 12 January 2023

Available online 20 January 2023

#### Keywords:

ClinVar

Developmental disorder

Exome reanalysis

Rare diseases

### ABSTRACT

**Purpose:** Within the Solve-RD project (<https://solve-rd.eu/>), the European Reference Network for Intellectual disability, TeleHealth, Autism and Congenital Anomalies aimed to investigate whether a reanalysis of exomes from unsolved cases based on ClinVar annotations could establish additional diagnoses. We present the results of the “ClinVar low-hanging fruit” reanalysis, reasons for the failure of previous analyses, and lessons learned.

**Methods:** Data from the first 3576 exomes (1522 probands and 2054 relatives) collected from European Reference Network for Intellectual disability, TeleHealth, Autism and Congenital Anomalies was reanalyzed by the Solve-RD consortium by evaluating for the presence of single-nucleotide variant, and small insertions and deletions already reported as (likely) pathogenic in ClinVar. Variants were filtered according to frequency, genotype, and mode of inheritance and reinterpreted.

**Results:** We identified causal variants in 59 cases (3.9%), 50 of them also raised by other approaches and 9 leading to new diagnoses, highlighting interpretation challenges: variants in genes not known to be involved in human disease at the time of the first analysis, misleading genotypes, or variants undetected by local pipelines (variants in off-target regions, low quality filters, low allelic balance, or high frequency).

**Conclusion:** The “ClinVar low-hanging fruit” analysis represents an effective, fast, and easy approach to recover causal variants from exome sequencing data, herewith contributing to the reduction of the diagnostic deadlock.

© 2023 The Authors. Published by Elsevier Inc. on behalf of American College of Medical Genetics and Genomics. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Introduction

Exome sequencing (ES) and genome sequencing (GS) are gold standard tests for the diagnosis of genetic developmental anomalies. Wright et al<sup>1</sup> reported an exome diagnostic rate of 42% in pediatric population for intellectual disability (ID). Overall, in developmental disorders, ES allows positive results in 25% to 45% of the individuals.<sup>1-3</sup>

Unfortunately, the remaining 55% to 75% of the cases are still left without a diagnosis.

Most disease-causing variants are located within the coding part of the genome,<sup>4</sup> and therefore, most cases could theoretically be solved by an exome analysis. In practice, however, the tools and knowledge available are sometimes still insufficient or incomplete at the time of initial analysis. Reanalysis of existing sequencing data offers the possibility

\*Correspondence and requests for materials should be addressed to Anne-Sophie Denomme-Pichon, Unité Fonctionnelle Innovation en Diagnostic génomique des maladies rares, FHU-TRANSLAD, CHU Dijon Bourgogne, 15 boulevard Maréchal de Lattre de Tassigny, 21070 Dijon, France. *E-mail address:* [anne-sophie.denomme-pichon@u-bourgogne.fr](mailto:anne-sophie.denomme-pichon@u-bourgogne.fr) OR Laurence Faivre, FHU-TRANSLAD, Centre de Génétique, Hôpital d'Enfants, CHU Dijon Bourgogne, 14 rue Paul Gaffarel, 21073 Dijon, France. *E-mail address:* [laurence.faivre@chu-dijon.fr](mailto:laurence.faivre@chu-dijon.fr)

A full list of authors and affiliations appears at the end of the paper.

doi: <https://doi.org/10.1016/j.gim.2023.100018>

1098-3600/© 2023 The Authors. Published by Elsevier Inc. on behalf of American College of Medical Genetics and Genomics. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

to increase the diagnostic yield in the light of improved bioinformatics pipelines and updated literature.<sup>5-21</sup> Data reanalysis or reevaluation can increase the diagnostic yield by 22% to 89%,<sup>8,12,16,21</sup> depending on the date of the initial analysis and on the time before the reanalysis. The American College of Medical Genetics and Genomics (ACMG) has published a series of points to consider regarding the reevaluation and reanalysis of genomic test results at various levels.<sup>22</sup> It recommends a periodic variant-level reevaluation and case-level reanalysis. In addition, for cases remaining unsolved, it is useful to keep updated phenotypic descriptions to improve the specificity of the phenotype, because this can help to increase the diagnostic yield as well. Teams often use a manual approach for routine reanalysis; however this is not sustainable at scale, especially for iterative reanalyses with unsolved cases accumulating over time. The bottleneck here is that these procedures are time-consuming and require dedicated personnel to be completed.

Solve-RD is a Horizon 2020-supported EU study bringing together more than 300 clinicians, scientists, and patient representatives from 51 sites in over 15 European countries. The project aims to establish a diagnosis for genetically undiagnosed individuals. The use of different standardized and automated bioinformatics pipelines makes systematic reinterpretation less labor-intensive on a large data set. As part of the Solve-RD project, it is expected that a data set composed of more than 19,000 exomes and genomes of genetically unsolved cases will be reanalyzed in several batches over 5 years.<sup>23</sup> Data are provided by centers associated with 4 core European Reference Networks (ERNs), including the ERN-ITHACA (Intellectual disability, TeleHealth, Autism and Congenital Anomalies). To facilitate such large-scale reanalysis, multiple working groups for data analysis have been established, each focusing on a specific type of variant (eg, copy number variants [CNVs], mitochondrial variants, de novo variants, and mobile element insertions).

The working group focused on the detection of single-nucleotide variants (SNVs) and small insertions and deletions (indels) of less than 40 base pairs (Solve-RD SNV-indel Working Group) aims to initially programmatically identify SNVs or indels already annotated in ClinVar as pathogenic or likely pathogenic (P/LP) variant: the so-called “low-hanging fruit” variants.<sup>19,24</sup> ClinVar is a freely accessible public database aggregating information about genomic variations and their relationship to diseases. It features more than 1.9 million recorded submissions<sup>25</sup> and contains, in particular, variants reported as P/LP by the genetics community.

In this article, we present the results of this first focused reanalysis, referred to as “ClinVar low-hanging fruit,” on the first batch of the data set composed of ERN-ITHACA negative ES and report the causes underlying missed diagnoses at first analysis.

## Materials and Methods

The steps of the study process are shown in [Figure 1](#).

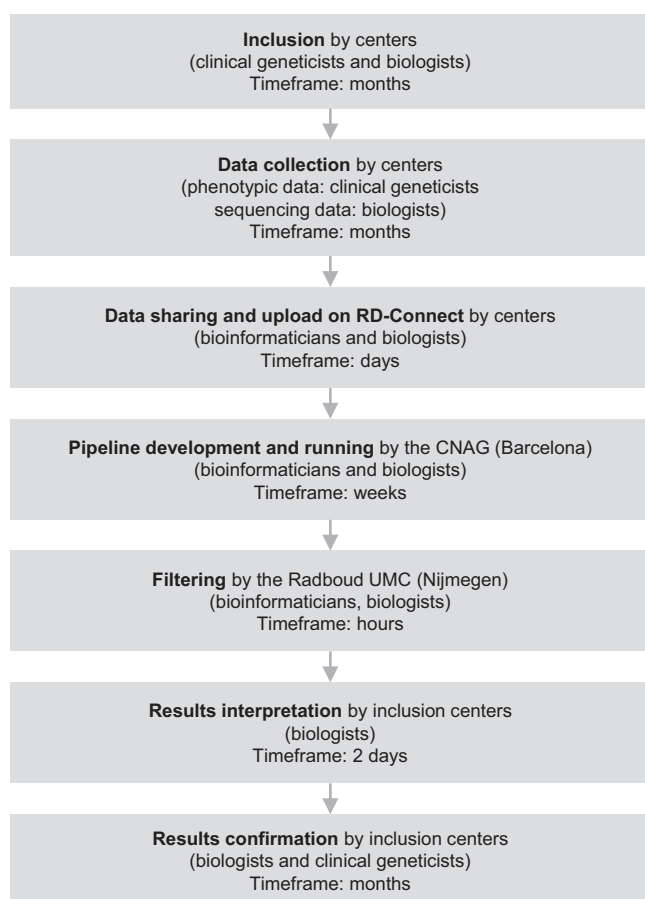
### Patient inclusion

Clinical data, pedigree structure, and their corresponding ES or GS raw data (FASTQ, BAM, or CRAM format) of 3576 data sets (including 1522 from index cases, with ID or developmental disorder, and 2054 relatives) have been shared internationally by 7 health care partners from ERN-ITHACA through the RD-Connect Genome-Phenome Analysis Platform (GPAP) (<https://platform.rd-connect.eu/>) between April 2018 and November 2019 as described by Zurek et al.<sup>23</sup> In compliance with the local ethical guidelines and the Declaration of Helsinki, all individuals (or legal representatives) provided informed consent to participate in the Solve-RD project.

### ES reanalysis and reinterpretation

ES and bioinformatics analyses were performed on platforms as previously described<sup>26-28</sup> and are detailed in the [Supplemental Methods](#). Raw data were reanalyzed using a centralized, automated analysis and filtering approach developed within the RD-Connect GPAP and in the context of the Solve-RD project.<sup>19</sup> We retained SNVs and indels (1) located in genes associated with ID or neurodevelopmental diseases, according to a list of clinical and research candidate genes provided by ERN-ITHACA (in [Supplemental Gene List Table](#)), (2) having a Genome Aggregation Database (gnomAD) allele frequency of < 1%, (3) having an internal RD-Connect GPAP allele frequency of < 2%, and (4) reported in ClinVar (v.13-01-2020)<sup>29</sup> as being likely pathogenic or pathogenic. Methods of this step have already been described in Matalonga et al.<sup>19</sup> Variants were prioritized if zygosity of the variant matched with the inheritance patterns of the gene (based on OMIM and the Developmental Disorder Gene-to-Phenotype [DDG2P] database) in which the variants were observed and, in addition, filtered on gnomAD frequency and internal frequency depending on the mode of inheritance ([Figure 2](#)). Frequency filters allow discarding frequent variants that are not involved in rare diseases and then avoid having to interpret irrelevant variants, including variants wrongly reported as being P/LP in ClinVar. The annotated variants were subsequently returned in February 2020 to the original submitting center that had initially produced the sequencing data for clinical interpretation.

Variants that were considered to be responsible for the phenotype of the cases were visualized in the Integrative Genomics Viewer (IGV) alongside the exome data of the parents when available for final visual valuation of quality controls (QCs). The ACMG and Association for Molecular Pathology criteria<sup>29</sup> for variant classification were used. Validation



**Figure 1** Steps, roles, and timeframes of the study. CNAG, National Center for Genomic Analysis; UMC, University Medical Center.

through alternative methods was performed locally in the clinical laboratory that submitted the sample to Solve-RD.

For each diagnosis achieved through this process, we subsequently attempted to determine the reason why the variants were not identified or discarded at the first analysis. Clinical laboratories, in parallel to submission of the genetically unsolved cases to Solve-RD, also performed local reanalysis. Only information about diagnoses obtained through this reanalysis strategy had to be recorded by the inclusion centers.

## Results

We identified 1618 SNVs and indels reported as being P/LP in ClinVar (Table 1, Figure 2) in 980 index cases from 649 trios, 11 duos, and 320 singletons. We identified 147 candidate variants in 127 individuals. Of these, the variant led to a conclusive molecular diagnosis in 59 individuals (3.9%). Among these 59 cases, 50 (3.3%) were also solved through local reanalyses in parallel to the Solve-RD project between the data upload and the reanalysis, in diagnostic or research laboratories, whereas the remaining 9 (0.6%) were solely identified

using the Solve-RD infrastructure. For various reasons, the 88 remaining variants did not solve the cases: single heterozygous variant for autosomal recessive disorders, no phenotypic fit combined with poor evidence of the pathogenicity based on the ACMG/Association for Molecular Pathology criteria<sup>29</sup> about the variants that should not have been reported in ClinVar as P/LP, variants already returned to the patient but not fully explanatory, or incidental findings. These variants were rejected based on expert interpretation in the context of the observed phenotype. We had no dual diagnoses.

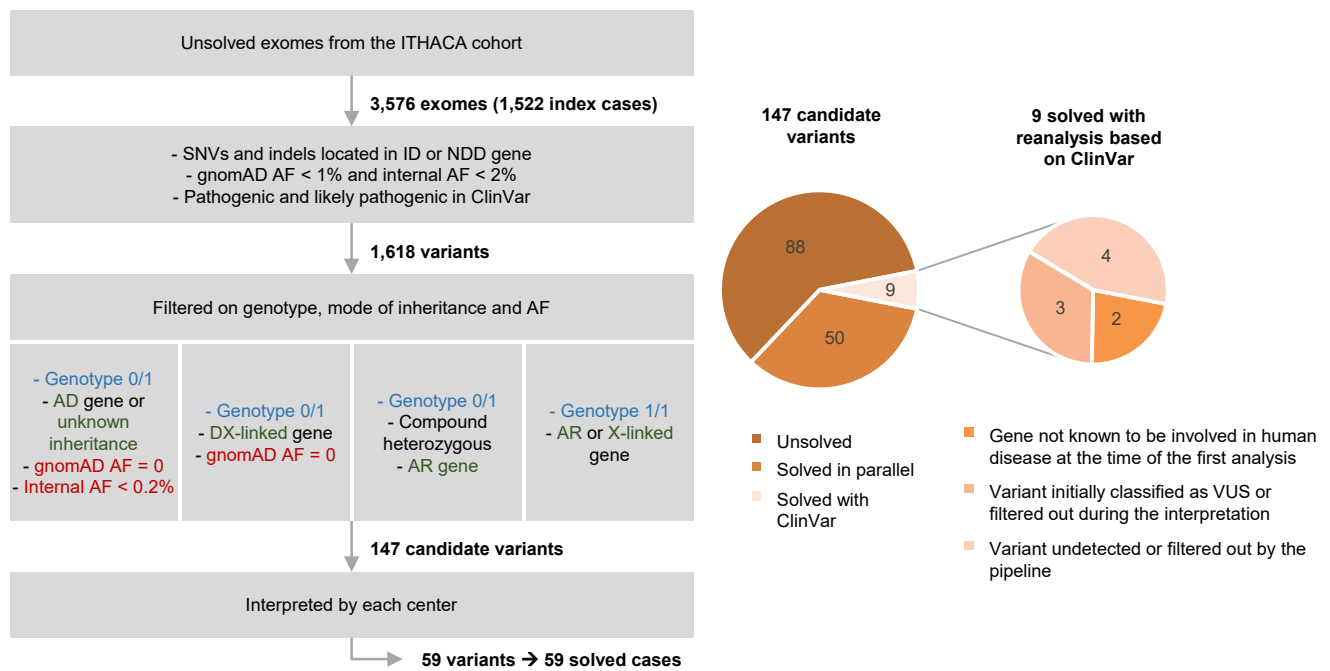
For the 9 cases only identified through the ClinVar Solve-RD analysis (Table 2), we retrospectively looked back at the original ES data and interpretations to understand why the variant was not considered at the time of the previous analyses. Two cases were not resolved, because the gene was not yet known to be involved in human disease (cases 1 and 2), 3 because the variants had been filtered out during data interpretation (cases 3-5) and 4 because the variants had not been detected using the in-house pipeline or were filtered out during the bioinformatics filtering process (cases 6-9).

### Case 1: *TRRAP*: No disease–gene association at the time of previous analyses

The first solved case was of an 8-year-old boy with intra-uterine growth retardation, global developmental delay, severe hypotonia, facial features (cleft palate, short upper lip), cerebellar hypoplasia, polymicrogyria, and an arachnoid cyst. The reanalysis identified a heterozygous missense variant in *TRRAP* reported as likely pathogenic in ClinVar, leading to the diagnosis of autosomal dominant developmental delay with or without dysmorphic facies and autism (OMIM 618454). At the time of the first exome analysis in 2018 using a singleton strategy, the *TRRAP* gene was not yet known to be involved in human disease. It has been published in 2019.<sup>30</sup> After the identification of the variant as an excellent candidate for the condition, family segregation using Sanger sequencing confirmed its de novo origin.

### Case 2: *NFIA*: No disease–gene association at the time of previous analyses

The second solved case was of a 12-year-old girl with an infantile-onset global developmental delay associated with behavioral abnormality, hypotonia, and corpus callosum hypoplasia. Dysmorphological features included epicanthal fold, strabismus, wide nasal base, brachydactyly and clinodactyly of the second toes, 1 hemangioma on the inner side of the thigh, and thickened skin. At 4 years and 4 months, she weighed 14.0 kg (−1.3 SD), she was 99.5 cm (−1.3 SD), and her occipital frontal circumference was 79.8 cm (−0.4 SD). The reanalysis identified a heterozygous missense variant in *NFIA* reported as likely pathogenic in ClinVar, associated with autosomal dominant brain malformations with or without urinary tract defects (OMIM



**Figure 2** Flowchart of the study. AD, autosomal disease; AF, allelic frequency; AR, autosomal recessive; NDD, neurodevelopmental disease; DX, dominant X-linked; gnomAD, Genome Aggregation Database; ID, intellectual disability; ITHACA, Intellectual disability, TeleHealth, Autism and Congenital Anomalies; SNV, single-nucleotide variant; VUS, variant of uncertain significance.

613735).<sup>31</sup> The variant was not identified at the time of the first exome analysis using a singleton strategy in 2016, because the *NFIA* gene was not yet reported in the OMIM database as a morbid gene. It was associated with human disease in the OMIM database in 2017. Family segregation using Sanger sequencing showed that the variant occurred de novo.

### Case 3: *PTEN*: A variant of uncertain significance reclassified as pathogenic

The third solved case was of a 9-year-old boy with speech and language developmental delay, impaired social interactions and poor eye contact, progressive macrocephaly, and short corpus callosum. His father presented with unilateral convergent strabismus, testicular ectopia, excessive

sweating, and macrocephaly (62.5 cm, > +4 SD). His mother also presented with macrocephaly (61 cm, +4 SD). He was born at 38 weeks of gestation via a normal delivery after a normal pregnancy. Head circumference at birth was 35.0 cm (−1.5 SD). At age 3 years, he weighed 17.0 kg (+0.8 SD) and was 90.0 cm tall (−2.5 SD) and he later developed macrocephaly (> +2 SD) and obesity (body mass index, +3.6 SD). He had dysmorphological features including frontal bossing, telecanthus, esotropia associated with horizontal nystagmus, and one cafe-au-lait spot.

The singleton-based reanalysis identified a heterozygous missense variant in *PTEN* that had been previously reported as a variant of uncertain significance in the patient because it was inherited from the mother, who was initially assumed to be unaffected. The variant had been submitted as pathogenic in ClinVar in the meantime. The Solve-RD reanalysis led us to reexamine the mother, who was apparently asymptom-

**Table 1** Number of cases and identified variants by center

Solve-RD Center From ERN-ITHACA	Uploaded Data Sets	Index Cases	Candidate Variants	Solved Cases	Cases Solved Only via the ClinVar-Based Reanalysis
Radboud University Medical Center (Nijmegen)	2031	687	63	23	4
Dijon Bourgogne University Hospital (Dijon)	826	554	62	36	5
Charles University (Prague)	232	105	13	0	0
Azienda Ospedaliero-Universitaria Senese (Siena)	231	83	8	0	0
Ospedale Pediatrico Bambino Gesù (Rome)	109	44	1	0	0
Bordeaux University Hospital (Bordeaux)	117	39	0	0	0
Institute of Health Carlos III (Madrid)	30	10	0	0	0
Total	3576	1522	147	59	9

ERN-ITHACA, European Reference Network for Intellectual disability, TeleHealth, Autism and Congenital Anomalies.

**Table 2** Results for the 9 additional diagnoses because of the reanalysis based on ClinVar

Case	Seq. Strategy	First Analysis (or Previous Reanalysis) Date	Disease (OMIM Number) and Mode of Inheritance	Gene	Genotype	cDNA Variant	Amino Acid Change	Inh.	ClinVar Status (Variant ID)	Submission Date in ClinVar	Cause of Missed Diagnosis
1	Singleton	2018-03	Developmental delay with or without dysmorphic facies and autism (618454), AD	<i>TRRAP</i>	Het	NM_001375524.1: c.3127G>A	p.(Ala1043Thr)	D.n.	P (634849)	2019-06	Gene not yet reported morbid in OMIM
2	Singleton	2016-08	Brain malformations with or without urinary tract defects (613735), AD	<i>NFIA</i>	Het	NM_001134673.4: c.361C>T	p.(Arg121Cys)	D.n.	P (265253)	2017-03	Gene not yet reported morbid in OMIM
3	Singleton	2016-08	Macrocephaly/autism syndrome (605309), AD	<i>PTEN</i>	Het	NM_000314.8: c.1004G>A	p.(Arg335Gln)	Mat.	P (427580)	2017-05	Segregation initially considered not compatible, because the variant was inherited from the mother, initially suspected as being unaffected
4	Singleton	2017-11	Spinocerebellar ataxia, autosomal recessive 10 (613728), AR	<i>ANO10</i>	Hom	NM_018075.5: c.132dup	p.(Asp45Argfs*9)	Mat. Pat.	P/LP (162016)	2014-12	Interpreted as heterozygous
5	Trio	2016-10	Intellectual developmental disorder, autosomal dominant 5 (612621), AD	<i>SYNGAP1</i>	Het	NM_006772.3: c.2294+1G>A	p.?	D.n.	LP (41460)	2013-04	Missed during interpretation by the geneticist
6	Trio	2018-12	Cortical dysplasia, complex, with other brain malformations 1 (614039), AD	<i>TUBB3</i>	Het	NM_006086.4: c.763G>A	p.(Val255Ile)	D.n.	LP (372654)	2016-12	Not targeted by the enrichment kit

(continued)

**Table 2** Continued

Case	Seq. Strategy	First Analysis (or Previous Reanalysis) Date	Disease (OMIM Number) and Mode of Inheritance	Gene	Genotype	cDNA Variant	Amino Acid Change	Inh.	ClinVar Status (Variant ID)	Submission Date in ClinVar	Cause of Missed Diagnosis
7	Trio	2019-07	Cortical dysplasia, complex, with other brain malformations 6 (615771), AD	<i>TUBB</i>	Het	NM_178014.4: c.895A>G	p.(Met299Val)	D.n.	P (127189)	2014-05	Low variant quality
8	Trio	2019-07	Developmental and epileptic encephalopathy 33 (616409), AD	<i>EEF1A2</i>	Mosaic (23%)	NM_001098.5: c.271G>A	p.(Asp91Asn)	D.n.	P/LP (279803)	2019-01	Low allelic balance
9	Singleton	2019-01	Ehlers-Danlos syndrome, kyphoscoliotic type 2 (614557), AR	<i>FKBP14</i>	Hom	NM_017946.4: c.362dup	p.(Glu122Argfs*7)	Mat. Pat.	P/LP (279809)	2012-03	High frequency in population

AD, autosomal dominant; AR, autosomal recessive; cDNA, complementary DNA; D.n, de novo; Het, heterozygous; Hom, homozygous; ID, identifier; Inh, inheritance; LP, likely pathogenic; Mat, maternal; P, pathogenic; Pat, paternal; Seq, sequencing.

atic, with mild features compatible with the diagnosis. This led to the diagnosis of autosomal dominant macrocephaly/autism syndrome (OMIM 605309), a condition that is associated with *PTEN* loss of function.

#### Case 4: *ANO10*: Homozygous variant rejected because falsely presumed to be heterozygous

The fourth solved case was of a 41-year-old female with a young adult-onset spinocerebellar ataxia, associated with nystagmus and hypoplasia of the cerebellar vermis. There was no medical family history. The first symptoms started at age 29 years with progressive dysarthria.

The singleton-based reanalysis identified a homozygous indel in *ANO10* reported 4 times as pathogenic and once as likely pathogenic in ClinVar at the moment of our reanalysis and already described in several studies.<sup>32</sup> At the time of first analysis, the variant was examined; however, it was previously wrongly interpreted by the geneticist as heterozygous based on inspection of the data in IGV. The genome aligner had produced alignment artifacts, effectively hiding the indel variant because of its position at the end of reads (Figure 3). The pipeline used HaplotypeCaller and thus did not require realignment around the indels. Because there is no local realignment around the indels on raw data, the BAM files loaded into IGV to visualize the variant did not show the indel correctly, suggesting that the variation was in a heterozygous state.

The new calling was suggestive of a homozygous variant, which we confirmed through Sanger sequencing verifying the homozygous state, and leading to the diagnosis of autosomal recessive spinocerebellar ataxia type 10 (OMIM 613728).

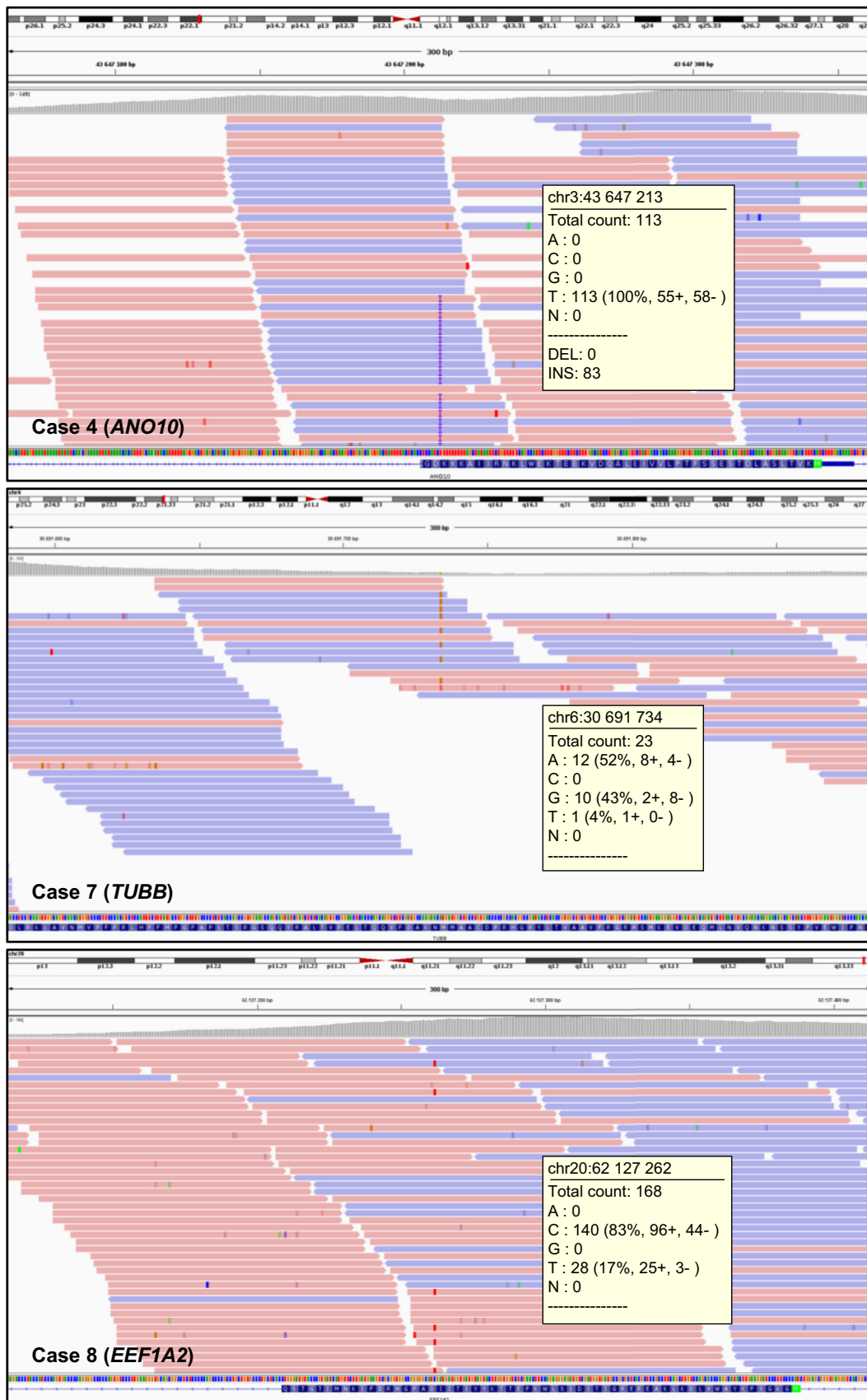
#### Case 5: *SYNGAP1*: Misinterpretation

The fifth solved case was of a 12-year-old female with ID, delayed speech and language development, absence seizures triggered by food intake, autism, and behavioral troubles. She had hyperbilirubinemia during neonatal period and feeding difficulties in infancy. She presented with hypopigmentation of the skin, recurrent infections, muscular hypotonia, hip dysplasia, and equinovarus deformity.

The reanalysis identified a de novo heterozygous splicing variant in *SYNGAP1*, reported as likely pathogenic in ClinVar, leading to the diagnosis of autosomal dominant intellectual developmental disorder type 5 (OMIM 612621). The variant was missed by the geneticist at the time of the trio-based analysis (2016), despite *SYNGAP1* being a known ID gene at the time of initial analysis, the variant being reported in the literature<sup>33</sup> and being clearly visible in the BAM file.

#### Case 6: *TUBB3*: Undetected by the pipeline

The sixth solved case has been previously described in de Boer et al.<sup>34</sup> This case was of a 16-year-old female with



**Figure 3** Integrative Genomics Viewer visualization of BAM files (GRCh37) from the first analysis of cases 4 (*ANO10*), 7 (*TUBB*), and 8 (*EEF1A2*). Case 4: chr3:g.43647213dup variant in the homozygous state (ref: 30; alt: 83). Case 7: chr6:g.30691734A>G variant with low depth of coverage (23 X: ref: 12; alt 11 with 10 G and 1 T). The variant is absent from the parents. Case 8: chr20:g.62127262C>T variant with low allelic balance (0.17%). alt, alteration; ref, reference.

severe neurodevelopmental delay, severe ID, and progressive microcephaly ( $-2$  SD at 1 year). Family history was negative for genetic diseases and ID. She presented with delayed motor and communication development and behavioral problems (auto- and hetero-aggressive behavior, sleep disturbance, and phonophobia). She had mild dysmorphological features, including round face, deeply set eyes, short philtrum, inverted nipples, pectus carinatum, and short feet. She had other medical problems, including short stature ( $-2.5$  SD), high hypermetropia, strabismus, recurrent upper respiratory and urinary tract infections, constipation, and impaired pain sensation. Brain magnetic resonance imaging (MRI) at age 2 years and 2 months showed notably corpus callosum hypoplasia, reduced volume of supratentorial white matter, and ventriculomegaly.

The reanalysis identified a heterozygous missense variant in *TUBB3*, reported as likely pathogenic in ClinVar. The in-house pipeline in the first trio-based analysis (2014) did not call the variant, because it was located in a region not targeted by the enrichment kit. Variant calling in the bioinformatics pipeline was based on this target set  $\pm 200$  bases, leaving the variant undetected. We found it in Solve-RD using genome-wide variant calling on raw data sets, leading to the diagnosis of autosomal dominant cortical dysplasia, complex, with other brain malformations type 1 (OMIM 614039).

### Case 7: *TUBB*: Rejected by the pipeline based on QC filters

The seventh solved case was of a 24-year-old female with global developmental delay (motor, speech, and language development) and cerebral palsy. She had severe cortical visual impairment associated with megalocornea, hypoplasia of the optic nerve, severe myopia, unilateral strabismus, horizontal nystagmus, and photophobia. Dysmorphological features included ptosis, wide nasal bridge, bulbous nasal tip, low insertion of columella, wide mouth, thick lower lip vermilion, long fingers with prominent fingertip pads, and short feet with pes planus and hallux valgus. Other medical problems included constipation and Hirschsprung disease. Brain MRI showed corpus callosum agenesis, gray matter heterotopias, and abnormality of the caudate nucleus and putamen. Mother presented with (familial occurring) bilateral coloboma of the iris and retina.

The reanalysis identified a heterozygous pathogenic missense in *TUBB* reported in the literature in 2012.<sup>35</sup> The variant had not passed at initial trio-based analysis in 2013 because of the overall low quality of the variant, likely because of low depth of coverage (23 X coverage, with 8 variant reads) (Figure 3). Indeed, the in-house pipeline generates a QC score per variant, and this variant had a low score, probably because of the low coverage and because there was 1 read with a third allele at the locus. The initial local diagnostic analysis heavily relied on de novo variant

calling, which is less accurate for variants with low QC scores. Therefore, the variant was likely filtered out during the interpretation process. Now, the variant led to the diagnosis of autosomal dominant cortical dysplasia, complex, with other brain malformations type 6 (OMIM 615771).

### Case 8: *EEF1A2*: Rejected by the pipeline based on allelic balance

The eighth solved case was of a 19-year-old male with severe ID associated with global developmental delay (motor, speech, and language development delay), autism (stereotypies, short attention span, poor eye contact), seizures, hypotonia, sleep disturbance, and gait troubles. Dysmorphological features included macrocephaly, retrognathia, high forehead, anteverted nares, macrotia, incisor macrodontia, large hands with abnormality of the thumbs, and hip dysplasia. Brain MRI showed nonspecific white matter abnormalities around the left frontal horn.

The reanalysis identified a pathogenic missense variant in *EEF1A2*, in a mosaic state in the blood (17%). At the time of the first trio-based analysis in 2014, the position of this variant was properly covered and the variant was identified with substandard QC, however was presumably subsequently rejected based on the allelic fraction being below the threshold for interpretation ( $<20\%$ ) (Figure 3). The variant led to a diagnosis of developmental and epileptic encephalopathy type 33 (OMIM 616409).

### Case 9: *FKBP14*: Filtered by pipeline as too frequent

The ninth solved case was of a 29-year-old female with hyperelastic skin, hypermobility, severe muscle hypotonia at birth, and delayed gross motor development. Family history was negative for genetic diseases. Her father was 170 cm ( $-1$  SD) and mother 160 cm ( $-1$  SD). She was born prematurely after a normal pregnancy. At 24 years old, she weighed 52.0 kg ( $-0.7$  SD), was 147 cm tall ( $-2.5$  SD) and her sitting height was 75.0 cm ( $-4.9$  SD). Dysmorphological features included progressive juvenile onset scoliosis, arachnodactyly, thin fibula, pes valgus, short fourth and fifth metatarsal, retrognathia, high narrow palate, and microtia. She presented with convergent strabismus and high progressive myopia. She was initially suspected of having Stickler syndrome.

The reanalysis identified a homozygous pathogenic frameshift in *FKBP14*. At the time of the first singleton-based analysis, the variant was present in the raw VCF, but it had been filtered out by the in-house pipeline because it was tagged as “common” in the single nucleotide polymorphism database (dbSNP build 151) with an allele frequency in gnomAD of over  $10^{-4}$ . It appears that this very frequent variant accounts for 70% of alleles involved in the



autosomal recessive Ehlers-Danlos syndrome, kyphoscoliotic type 2 (OMIM 614557).<sup>36</sup>

## Discussion

The aim of Solve-RD is to solve the diagnostic dead end via reanalysis and also to identify new genes and new molecular mechanisms through data sharing and patient inclusion at a pan-European level. The purpose of the “ClinVar low-hanging fruit” analysis of 3576 individuals (including 1522 index cases) from the ERN-ITHACA cohort was not to identify new genes but to decrease the diagnostic dead end. With 59 diagnoses (3.9%), it shows the usefulness of a systematic reanalysis of negative exomes that is focused on the SNVs and indels reported as being pathogenic and likely pathogenic in ClinVar through systematic realignment, recalling, and reinterpretation using up-to-date bioinformatics pipelines.<sup>19</sup> Between the inclusion of the cases in the Solve-RD project in 2018 and the reanalysis in 2020, 50 candidate cases from the “ClinVar low-hanging fruit” reanalysis were solved in parallel to the Solve-RD project, mainly through co-occurring reanalyses or re-evaluation of the patients (because individuals without a genetic diagnosis are often advised to recontact their clinician for such analysis). Nine additional cases (0.6%) were not yet identified locally either because they had not recontacted their clinician for reanalysis or if they had, because the variant escaped detection at any stage of the ES process (enrichment, efficient sequencing, alignment, calling, annotation, and/or interpretation). These results match those of the previous similar study on the Deciphering Developmental Disorders cohort (13,462 probands) in which Wright et al<sup>37</sup> were able to identify 112 variants in 107 probands (0.8%) as possible diagnoses.

After collection of the raw data and completion of their bioinformatics reanalysis, the reinterpretation of the 1522 negative cases was easy to implement and fast to interpret. There were only 147 variants after overlap with a disease-specific gene list and prioritization. Although this list limited the possibility of identifying so-called unanticipated diagnosis, referring to phenocopies of a genetically different disease group (eg, ID genes vs epilepsy genes), this approach does help to limit the risks of incidental findings. In addition, the *in silico* enrichment filters also limited the needs in human resources, because the centers did not have to filter the thousands of cases manually. Moreover, the cohort analysis with only candidate variants to re-evaluate resulted in time saving compared with a case-by-case reanalysis considering different strategies.

Our centralized reanalysis has taught us some lessons. The first and very essential one is that stringent filters should not be used for variants reported as P/LP in ClinVar: (1) if there is a variant with low depth of coverage, consider it nonetheless for the interpretation, (2) if there is a variant with low allelic balance, consider mosaicism, (3) if there is

an inherited variant, check for paucisymptomatic parent or consider incomplete penetrance, and (4) if there is a frequent variant, check for recurrent variants and involvement in recessive disorder and use homozygous count annotation in gnomAD. The second lesson is to remain cautious when interpreting variants located at the end of the reads, especially in IGV. Genome aligners can produce alignment artifacts and hide indels located at the end of the reads, especially when there is no realignment around the indel step in the bioinformatic pipeline, because it can be the case when using UnifiedGenotyper. In these cases, the variant calling is correct, but the BAM still contains the misalignment. Although most current pipelines use variant callers with a reassembly step, such as HaplotypeCaller or Platypus, and thus do not require local indel realignment, it is still useful for legacy tools such as UnifiedGenotyper to correct mapping errors made by genome aligners.<sup>38,39</sup> We have to keep in mind that when looking at variations in IGV located at the end of reads, there may be artifacts related to the fact that there was no realignment around the indels. The third lesson is to never stop reanalyzing genomic data because new genes and new variants are reported all the time. Some tools allow automating the reanalysis in real time,<sup>13</sup> such as Variant Alert,<sup>40</sup> which can highlight updates from ClinVar as soon as variants are submitted at the scale of the complete database.

ClinVar-based reanalysis is an effective technique for solving unsolved cases from the ERN-ITHACA cohort with ID and congenital abnormalities, and this strategy should also be effective for other types of cohorts such as those of Solve-RD project. Given the success of the ClinVar-based reanalysis in diagnosing unsolved cases, similar strategies for other variant types, such as CNVs, may be equally successful, which are currently ongoing.

This approach should be used as part of a more comprehensive reanalysis strategy, because it only targets variants already reported and cannot resolve all cases. Other strategies should be considered in reanalysis, such as the identification of *de novo* variants from trio, CNVs, mitochondrial variants, mobile element insertions, short tandem repeat expansions, uniparental disomy, and variants in runs of homozygosity. However, ES reanalysis alone will not solve the other causes of diagnostic impasse and additional experiments such as GS or multiomic analyses will be therefore required.

## Data Availability

Genomic and phenotypic data used in this study are available on the RD-Connect Genome-Phenome Analysis Platform, accessible to researchers and clinicians (<https://platform.rd-connect.eu/userregistration/>). The RD-Connect identifiers of the cases of this cohort are available from the corresponding author on request.

## Acknowledgments

We are very grateful to the families participating in the Solve-RD project. We thank the computer cluster of the University of Burgundy for technical support and management of the Dijon Bourgogne computing platform.

## Funding

The Solve-RD project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 779257. Data were analyzed using the RD-Connect Genome-Phenome Analysis Platform, which received funding from the EU projects RD-Connect, Solve-RD, and European Joint Programme on Rare Diseases (grant numbers FP7 305444, H2020 779257, H2020 825575), Instituto de Salud Carlos III (grant numbers PT13/0001/0044, PT17/0009/0019; Instituto Nacional de Bioinformática), and ELIXIR Implementation Studies. The collaborations in this study were facilitated by the European Reference Network for Intellectual disability, TeleHealth, Autism and Congenital Anomalies, one of the 24 European Reference Networks approved by the European Reference Network Board of Member States, cofunded by the European Commission. This project was supported by the Czech Ministry of Health (number 00064203) and by the Czech Ministry of Education, Youth and Sports (number - LM2018132) to M.M.

## Author Information

Data Curation: A.-S.D.-P., E.d.B., A.C., N.J.-M., S.M., S.P., F.C.R., M.S., M.T.; Formal Analysis: A.-S.D.-P., E.d.B., T.B.H., C.T.-R., F.T.M.-T., C.P.; Funding Acquisition: H.G., B.Z.; Investigation: A.-S.D.-P., A.T.; Project Administration: K.E., H.G., B.Z.; Software: L.M., A.H.; Supervision: B.D., M.T., L.E.L.M.V., L.F.; Writing-original draft: A.-S.D.-P.; Writing-review and editing: A.-S.D.-P., L.M., E.d.B., E.B., A.C., B.D., C.F., C.G., H.G., T.B.H., E.L.-M., M.A.M., S.P., M.P., F.C.R., A.R., M.T., C.T.-R., F.T.M.-T., B.B.A.d.V., C.P., A.Vi., L.E.L.M.V., L.F.

## Ethics Declaration

This study was approved by the appropriate French independent ethics committee (Comité de Protection des Personnes Ouest I 2019T2-12). Individuals from Prague were included in the collection from the Coordinating Center for Patients with Rare Diseases (number NCCVO 135V32R001102, 21.07.2015). Individuals from Nijmegen

were included in the Solve-RD collection (identifier [ID]: 2018-4986, 28.11.2018; Biobank contract between department and Radboudumc Biobank ID: 2014-1254, 2014-09-09). Individuals from Roma were included through the European Commission's information system for patient sharing as part of what is provided for European Reference Networks for Rare Diseases (2017-11-22). Individuals from Siena were included in the NeuroNext collection (Identificazione di geni coinvolti nella patogenesi dei disturbi del neurosviluppo e delle malattie neurodegenerative, 2020-05-18). Individuals from Madrid provided informed consent for the donation of biological samples to the National Rare Diseases Biobank of the Instituto de Investigación de enfermedades raras V.1 (ID: CEI PI 74\_2016, 2017-02-06).


## Conflict of Interest

The authors declare no conflicts of interest.

## Additional Information

The online version of this article (<https://doi.org/10.1016/j.gim.2023.100018>) contains supplementary material, which is available to authorized users.

## Authors

Anne-Sophie Denommé-Pichon<sup>1,2,\*</sup> , Leslie Matalonga<sup>3</sup>, Elke de Boer<sup>4,5</sup>, Adam Jackson<sup>6</sup>, Elisa Benetti<sup>7</sup>, Siddharth Banka<sup>6</sup>, Ange-Line Bruel<sup>1,2</sup>, Andrea Ciolfi<sup>8</sup>, Jill Clayton-Smith<sup>6</sup>, Bruno Dallapiccola<sup>8</sup>, Yannis Duffourd<sup>1,2</sup>, Kornelia Ellwanger<sup>9,10</sup>, Chiara Fallerini<sup>7,11</sup>, Christian Gilissen<sup>4,12</sup>, Holm Graessner<sup>9,10</sup>, Tobias B. Haack<sup>9,10</sup>, Marketa Havlovicova<sup>13</sup>, Alexander Hoischen<sup>4,12,14</sup>, Nolwenn Jean-Marçais<sup>2,15</sup>, Tjitske Kleefstra<sup>4,5,16</sup>, Estrella López-Martín<sup>17</sup>, Milan Macek Jr<sup>13</sup>, Maria Antonietta Mencarelli<sup>18</sup>, Sébastien Moutton<sup>2</sup>, Rolph Pfundt<sup>4,5</sup>, Simone Pizzi<sup>8</sup>, Manuel Posada<sup>17</sup>, Francesca Clementina Radio<sup>8</sup>, Alessandra Renieri<sup>7,11,18</sup>, Caroline Rooryck<sup>19</sup>, Lukas Ryba<sup>13</sup>, Hana Safradou<sup>1,2</sup>, Martin Schwarz<sup>13</sup>, Marco Tartaglia<sup>8</sup>, Christel Thauvin-Robinet<sup>1,2,15</sup>, Julien Thevenon<sup>2</sup>, Frédéric Tran Mau-Them<sup>1,2</sup>, Aurélien Trimouille<sup>20</sup>, Pavel Votycka<sup>13</sup>, Bert B.A. de Vries<sup>4,5</sup>, Marjolein H. Willemsen<sup>4,5</sup>, Birte Zurek<sup>9,10</sup>, Alain Verloes<sup>21,22</sup>, Christophe Philippe<sup>1,2</sup>, Solve-RD DITF-ITHACA, Solve-RD SNV-indel Working Group, Solve-RD Consortia, Orphanomix Group, Antonio Vitobello<sup>1,2</sup>, Lisenka E.L.M. Vissers<sup>4,5</sup>, Laurence Faivre<sup>2,15,\*</sup>

## Affiliations

<sup>1</sup>Functional Unit for Diagnostic Innovation in Rare Diseases, FHU-TRANSLAD, Dijon Bourgogne University Hospital, Dijon, France; <sup>2</sup>INSERM UMR1231 GAD “Génétique des Anomalies du Développement,” FHU-TRANSLAD, University of Burgundy, Dijon, France; <sup>3</sup>CNAG-CRG, Centre for Genomic Regulation,” The Barcelona Institute of Science and Technology, Barcelona, Spain; <sup>4</sup>Department of Human Genetics, Radboudumc, Nijmegen, The Netherlands; <sup>5</sup>Donders Institute for Brain, Cognition and Behaviour, Radboudumc, Nijmegen, The Netherlands; <sup>6</sup>Manchester Centre for Genomic Medicine, University of Manchester, Manchester, United Kingdom; <sup>7</sup>MedBiotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, Siena, Italy; <sup>8</sup>Molecular Genetics and Functional Genomics, Ospedale Pediatrico Bambino Gesù, IRCCS, Rome, Italy; <sup>9</sup>Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany; <sup>10</sup>Centre for Rare Diseases, University Hospital Tübingen, Tübingen, Germany; <sup>11</sup>Medical Genetics, University of Siena, Siena, Italy; <sup>12</sup>Radboud Institute for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands; <sup>13</sup>Department of Biology and Medical Genetics, Second Faculty of Medicine of Charles University and Motol University Hospital, Prague, Czech Republic; <sup>14</sup>Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboudumc, Nijmegen, The Netherlands; <sup>15</sup>Department of Genetics and Reference Center for Development Disorders and Intellectual Disabilities, FHU-TRANSLAD and GIMI Institute, Dijon Bourgogne University Hospital, Dijon, France; <sup>16</sup>Center of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands; <sup>17</sup>Institute of Rare Diseases Research, Spanish Undiagnosed Rare Diseases Cases Program (SpainUDP) & Undiagnosed Diseases Network International, Instituto de Salud Carlos III, Madrid, Spain; <sup>18</sup>Medical Genetics, Azienda Ospedaliero-Universitaria Senese, Siena, Italy; <sup>19</sup>MRGM INSERM U1211, University of Bordeaux, Medical Genetics Department, Bordeaux University Hospital, Bordeaux, France; <sup>20</sup>Molecular Genetics Laboratory, Medical Genetics Department, Bordeaux University Hospital – Hôpital Pellegrin, Bordeaux, France; <sup>21</sup>Department of Genetics, Assistance Publique-Hôpitaux de Paris - Université de Paris, Paris, France; <sup>22</sup>INSERM UMR 1141 “NeuroDiderot,” Hôpital Robert Debré, Paris, France

## Members of Solve-RD DITF-ITHACA, Solve-RD SNV-indel Working Group, Solve-RD Consortia, and Orphanomix Group

### Solve-RD DITF-ITHACA

Kristin M. Abbott, Siddharth Banka, Elke de Boer, Andrea Ciolfi, Jill Clayton-Smith, Bruno Dallapiccola, Anne-Sophie

Denommé-Pichon, Laurence Faivre, Christian Gilissen, Tobias B. Haack, Marketa Havlovicova, Alexander Hoischen, Adam Jackson, Mieke Kerstjens, Tjitske Kleefstra, Estrella López Martín, Milan Macek Jr, Leslie Matalonga, Isabelle Maystadt, Manuela Morleo, Vincenzo Nigro, Michele Pinelli, Simone Pizzi, Manuel Posada, Francesca C. Radio, Alessandra Renieri, Olaf Riess, Caroline Rooryck, Lukas Ryba, Jean-Madeleine de Sainte Agathe, Gijs W.E. Santen, Martin Schwarz, Marco Tartaglia, Christel Thauvin, Annalaura Torella, Aurélien Trimouille, Alain Verloes, Lisenka Vissers, Antonio Vitobello, Pavel Votypka, and Kristina Zguro

### Solve-RD SNV-indel working group

Elke de Boer, Enzo Cohen, Daniel Danis, Anne-Sophie Denommé-Pichon, Fei Gao, Christian Gilissen, Rita Horvath, Mridul Johari, Lennart Johanson, Shuang Li, Leslie Matalonga, Heba Morsy, Isabelle Nelson, Ida Paramonov, Iris B.A.W. te Paske, Peter Robinson, Marco Savarese, Wouter Steyaert, Ana Töpf, Aurélien Trimouille, Joeri K. van der Velde, Jana Vandrovicova, Antonio Vitobello

### Solve-RD Consortia

#### The Eberhard Karls Universitaet Tuebingen (EKUT)

Olaf Riess, Tobias B. Haack, Holm Graessner, Birte Zurek, Kornelia Ellwanger, Stephan Ossowski, German Demidov, Marc Sturm, Julia M. Schulze-Hentrich, Rebecca Schüle, Jishu Xu, Christoph Kessler, Melanie Wayand, Matthias Synofzik, Carlo Wilke, Andreas Traschütz, Ludger Schöls, Holger Hengel, Holger Lerche, Josua Kegele, Peter Heutink

#### Radboud University Medical Center (RUMC)

Han Brunner, Hans Scheffer, Noline Hoogerbrugge, Alexander Hoischen, Peter A.C. 't Hoen, Lisenka E.L.M. Vissers, Christian Gilissen, Wouter Steyaert, Karolis Sablauskas, Richarda M. de Voer, Erik-Jan Kamsteeg, Bart van de Warrenburg, Nienke van Os, Iris te Paske, Erik Janssen, Elke de Boer, Marloes Steehouwer, Burcu Yaldiz, Tjitske Kleefstra

#### University of Leicester

Anthony J. Brookes, Colin Veal, Spencer Gibson, Vatsalya Maddi, Mehdi Mehtarizadeh, Umar Riaz, Greg Warren, Farid Yavari Dizjikan, Thomas Shorter

#### Newcastle University (UNEW)

Ana Töpf, Volker Straub, Chiara Marini Bettolo, Jordi Diaz Manera, Sophie Hambleton, Karin Engelhardt

#### Manchester University Hospital (MUH)

Jill Clayton-Smith, Siddharth Banka, Elizabeth Alexander, Adam Jackson

#### Dijon Bourgogne University Hospital

Laurence Faivre, Christel Thauvin, Antonio Vitobello, Anne-Sophie Denommé-Pichon, Yannis Duffourd, Angeline Bruel, Christine Peyron, Aurore Pélissier

**National Center for Genomic Analysis-Center for Genomic Regulation (CNAG-CRG)**

Sergi Beltran, Ivo Glynne Gut, Steven Laurie, Davide Piscia, Leslie Matalonga, Anastasios Papakonstantinou, Gemma Bullich, Alberto Corvo, Marcos Fernandez-Callejo, Carles Hernández, Daniel Picó, Ida Paramonov, Hanns Lochmüller

**EURORDIS**

Gulcin Gumus, Virginie Bros-Facer

**National Institute of Health and Medical Research (INSERM) - Orphanet**

Ana Rath, Marc Hanauer, David Lagorce, Oscar Hongnat, Maroua Chahdil, Emeline Lebreton

**National Institute of Health and Medical Research - Institute for Brain and Spinal Cord (INSERM-ICM)**

Giovanni Stevanin, Alexandra Durr, Claire-Sophie Davoine, Léna Guillot-Noel, Anna Heinzmann, Giulia Coarelli

**National Institute of Health and Medical Research - Institute of Myology (INSERM-CRM)**

Gisèle Bonne, Teresinha Evangelista, Valérie Allamand, Isabelle Nelson, Rabah Ben Yaou, Corinne Metay, Bruno Eymard, Enzo Cohen, Antonio Atalaia, Tanya Stojkovic

**Charles University**

Milan Macek Jr, Marek Turnovec, Dana Thomasová, Radka Pourová Kremlíková, Vera Franková, Markéta Havlovicová, Petra Lišková, Pavla Doležalová

**European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI)**

Helen Parkinson, Thomas Keane, Mallory Freeberg, Coline Thomas, Dylan Spalding

**Jackson Laboratory**

Peter Robinson, Daniel Danis

**King's College London (KCL)**

Glenn Robert, Alessia Costa, Christine Patch

**University College London - Queen Square Institute of Neurology (UCL-IoN)**

Mike Hanna, Henry Houlden, Mary Reilly, Jana Vandrovcova, Stephanie Efthymiou, Heba Morsy, Elisa Cali, Francesca Magrinelli, Sanjay M. Sisodiya, Jonathan Rohrer

**University College London - Institute of Child Health (UCL-ICH)**

Francesco Muntoni, Irina Zaharieva, Anna Sarkozy

**University of Antwerp**

Vincent Timmerman, Jonathan Baets, Geert de Vries, Jonathan De Winter, Danique Beijer, Peter de Jonghe,

Liedewei Van de Vondel, Willem De Ridder, Sarah Weckhuysen

**University of Naples/Telethon UnDiagnosed Disease (UDP)**

Vincenzo Nigro, Margherita Mutarelli, Manuela Morleo, Michele Pinelli, Alessandra Varavallo, Sandro Banfi, Annalaura Torella, Francesco Musacchia, Giulio Piluso

**University of Ferrara (UNIFE)**

Alessandra Ferlini, Rita Selvatici, Francesca Gualandi, Stefania Bigoni, Rachele Rossi, Marcella Neri

**University Hospital Bonn (UKB)**

Stefan Aretz, Isabel Spier, Anna Katharina Sommer, Sophia Peters

**Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP)**

Carla Oliveira, Jose Garcia Pelaez, Ana Rita Matos, Celina São José, Marta Ferreira, Irene Gullo, Susana Fernandes, Luzia Garrido, Pedro Ferreira, Fátima Carneiro

**University Medical Center Groningen (UMCG)**

Morris A. Swertz, Lennart Johansson, Joeri K. van der Velde, Gerben van der Vries, Pieter B. Neerinx, David Ruvolo, Kristin M. Abbott, Wilhemina S. Kerstjens Frederikse, Eveline Zonneveld-Huijssoon, Dieuwke Roelofs-Prins, Marielle van Gijn<sup>83,40</sup>

**Charité Berlin**

Sebastian Köhler

**Sheffield Hallam University (SHU)**

Alison Metcalfe

**Assistance publique - Hôpitaux de Paris (APHP)**

Alain Verloes, Séverine Drunat, Delphine Heron, Cyril Mignot, Boris Keren, Jean-Madeleine de Sainte Agathe

**Bordeaux University Hospital**

Caroline Rooryck, Didier Lacombe, Aurelien Trimouille

**Catalan Institute of Oncology (ICO)**

Gabriel Capella, Laura Valle

**Ludwig Maximilian University of Munich (KU Munich)**

Elke Holinski-Feder, Andreas Laner, Verena Steinke-Lange

**Cliniques universitaires Saint-Luc (CUSL)**

Maria-Roberta Cilio, Evelina Carpanca, Chantal Depondt, Damien Lederer, Yves Sznajer, Sarah Duerinckx, Sandrine Mary

**Vall d'Hebron University Hospital**

Alfons Macaya, Ana Cazorro-Gutiérrez, Belén Pérez-Dueñas, Francina Munell, Clara Franco Jarava, Laura Batlle Masó, Anna Marcé-Grau, Roger Colobran

**Folkhälsan Research Centre**

Peter Hackman, Mridul Johari, Marco Savarese, Bjarne Udd

**Ghent University Hospital**

Dimitri Hemelsoet, Bart Dermaut, Nika Schuermans, Bruce Poppe, Hannah Verdin

**Hospital Sant Joan de Déu Barcelona**

Andrés Nascimento Osorio

**Institute of Human Genetics, University Hospital Essen**

Christel Depienne, Andreas Roos

**Institute of Pathology and Genetics of Charleroi**

Isabelle Maystadt

**Technical University Munich**

Isabell Cordts, Marcus Deschauer

**L'Istituto di Ricovero e Cura a Carattere Scientifico  
Giannina Gaslini (IRCCS-IGG)**

Pasquale Striano, Federico Zara, Antonella Riva, Michele Iacomino, Paolo Uva, Marcello Scala, Paolo Scudieri

**Koç University**

Ayşe Nazlı Başak

**Katholieke Universiteit Leuven (KU Leuven)**

Kristl Claeys

**Ludwig Boltzmann Institute for Rare and Undiagnosed  
Diseases**

Kaan Boztug, Matthias Haimel

**Leiden University Medical Center (LUMC)**

Gijs W.E, Claudia A.L. Ruivenkamp, Daniel Natera de Benito

**University of Freiburg**

Hanns Lochmüller, Rachel Thompson, Kiran Polavarapu, Bodo Grimbacher

**Neurology/Neurogenetics Laboratory University of  
Crete, Heraklion**

Ioannis Zaganas, Evgenia Kokosali, Mathioudakis Lambros, Athanasios Evangeliou, Martha Spilioti, Elisabeth Kapaki, Mara Bourboulis

**Ospedale Pediatrico Bambino Gesù, Rome**

Andrea Ciolfi, Bruno Dallapiccola, Simone Pizzi, Francesca Clementina Radio, Marco Tartaglia

**Semmelweis University Budapest**

Peter Balicza, Maria Judit Molnar

**Spain UnDiagnosed Disease Program (UDP)**

Manuel Posada De la Paz, Eva Bermejo Sánchez, Estrella López Martín, Beatriz Martínez Delgado, F. Javier Alonso García de la Rosa

**Dresden University of Technology (TU Dresden)**

Evelin Schröck, Andreas Rump

**University Hospital Meyer, Florence**

Davide Mei, Annalisa Vetro, Simona Balestrini, Renzo Guerrini

**University of Cambridge**

Rita Horvath, Patrick F. Chinnery, Thiloka Ratnaike, Fei Gao, Katherine Schon

**University of Ljubljana**

Ales Maver, Borut Peterlin

**University of Lübeck**

Alexander Münchau, Katja Lohmann, Rebecca Herzog, Martje Pauly

**University of Luxembourg**

Patrick May

**University of Oxford**

David Beeson, Judith Cossins

**University of Siena**

Alessandra Renieri, Simone Furini, Chiara Fallerini, Elisa Benetti

**Orphanomix Group**

Alexandra Afenjar, Alice Goldenberg, Alice Masurel, Alice Phan, Anne Dieux-Coeslier, Anne Fargeot, Anne-Marie Guerrot, Annick Toutain, Arnaud Molin, Arthur Sorlin, Audrey Putoux, Béatrice Jouret, Béatrice Laudier, Bénédicte Demeer, Bérénice Doray, Bertille Bonniaud, Bertrand Isidor, Brigitte Gilbert-Dussardier, Bruno Leheup, Bruno Reversade, Carle Paul, Catherine Vincent-Delorme, Cecilia Neiva, Céline Poirsier, Chloé Quélin, Christine Chiaverini, Christine Coubes, Christine Francannet, Cindy Colson, Claire Desplantes, Constance Wells, Cyril Goizet, Damien Lederer, Damien Sanlaville, Daniel Amram, Daphné Lehalle, David Geneviève, Delphine Heron, Didier Lacombe, Dominique Gaillard, Einat Zivi, Elisabeth Sarrazin, Elisabeth Steichen, Élise Schaefer, Elodie Lacaze, Emmanuel Jacquemin, Ernie Bongers, Esra Kilic, Estelle Colin, Fabienne Giuliano, Fabienne Prieur, Fanny Laffargue, Fanny Morice-Picard, Florence Petit, François Cartault, François Feillet, Geneviève Baujat, Gilles Morin, Gwenaëlle Diene, Hubert Journal, Isabelle Maystadt, Isabelle Perthus, James Lespinasse, Jean-Luc Alessandri, Jeanne Amiel, Jelena Martinovic, Julian Delanne, Juliette

Albuisson, Laëtitia Lambert, Laurence Perrin, Lilian Bomme Ousager, Lionel Van Maldergem, Lucile Pinson, Lyse Ruaud, Mahtab Samimi, Marie Bournez, Marie Noëlle Bonnet-Dupeyron, Marie Vincent, Marie-Line Jacquemont, Marie-Pierre Cordier-Alex, Marion Gérard-Blanluet, Marjolaine Willems, Marta Spodenkiewicz, Martine Doco-Fenzy, Massimiliano Rossi, Mathilde Renaud, Mélanie Fradin, Michèle Mathieu, Muriel H. Holder-Espinasse, Nada Houcinat, Nadine Hanna, Nathalie Leperrier, Nicolas Chassaing, Nicole Philip, Odile Boute, Philippe Khau Van Kien, Philippe Parent, Pierre Bitoun, Pierre Sarda, Pierre Vabres, Pierre-Simon Jouk, Renaud Touraine, Salima El Chehadeh, Sandra Whalen, Sandrine Marlin, Sandrine Passemard, Sarah Grotto, Séverine Audebert Bellanger, Sophie Blesson, Sophie Nambot, Sophie Naudion, Stanislas Lyonnet, Sylvie Odent, Tania Attie-Bitach, Tiffany Busa, Valérie Drouin-Garraud, Valérie Layet, Varoona Bizaoui, Véronica Cusin, Yline Capri, Yves Alembik

## References

1. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet.* 2018;19(5):253-268. <http://doi.org/10.1038/nrg.2017.116>
2. Dong X, Liu B, Yang L, et al. Clinical exome sequencing as the first-tier test for diagnosing developmental disorders covering both CNV and SNV: a Chinese cohort. *J Med Genet.* 2020;57(8):558-566. <http://doi.org/10.1136/jmedgenet-2019-106377>
3. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet.* 2015;385(9975):1305-1314. [http://doi.org/10.1016/S0140-6736\(14\)61705-0](http://doi.org/10.1016/S0140-6736(14)61705-0)
4. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A.* 2009;106(45):19096-19101. <http://doi.org/10.1073/pnas.0910672106>
5. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med.* 2017;19(2):209-214. <http://doi.org/10.1038/gim.2016.88>
6. Al-Nabhani M, Al-Rashdi S, Al-Murshedi F, et al. Reanalysis of exome sequencing data of intellectual disability samples: yields and benefits. *Clin Genet.* 2018;94(6):495-501. <http://doi.org/10.1111/cge.13438>
7. Nambot S, Thevenon J, Kuentz P, et al. Clinical whole-exome sequencing for the diagnosis of rare disorders with congenital anomalies and/or intellectual disability: substantial interest of prospective annual reanalysis. *Genet Med.* 2018;20(6):645-654. <http://doi.org/10.1038/gim.2017.162>
8. Ewans LJ, Schofield D, Shrestha R, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med.* 2018;20(12):1564-1574. <http://doi.org/10.1038/gim.2018.39>
9. Genetics Initiative Epilepsy. The Epilepsy Genetics Initiative: systematic reanalysis of diagnostic exomes increases yield. *Epilepsia.* 2019;60(5):797-806. <http://doi.org/10.1111/epi.14698>
10. Al-Murshedi F, Meftah D, Scott P. Underdiagnoses resulting from variant misinterpretation: time for systematic reanalysis of whole exome data? *Eur J Med Genet.* 2019;62(1):39-43. <http://doi.org/10.1016/j.ejmg.2018.04.016>
11. Baker SW, Murrell JR, Nesbitt AI, et al. Automated clinical exome reanalysis reveals novel diagnoses. *J Mol Diagn.* 2019;21(1):38-48. <http://doi.org/10.1016/j.jmoldx.2018.07.008>
12. Li J, Gao K, Yan H, et al. Reanalysis of whole exome sequencing data in patients with epilepsy and intellectual disability/mental retardation. *Gene.* 2019;700:168-175. <http://doi.org/10.1016/j.gene.2019.03.037>
13. Liu P, Meng L, Normand EA, et al. Reanalysis of clinical exome sequencing data. *N Engl J Med.* 2019;380(25):2478-2480. <http://doi.org/10.1056/NEJMc1812033>
14. Schmitz-Abe K, Li Q, Rosen SM, et al. Unique bioinformatic approach and comprehensive reanalysis improve diagnostic yield of clinical exomes. *Eur J Hum Genet.* 2019;27(9):1398-1405. <http://doi.org/10.1038/s41431-019-0401-x>
15. Jalkh N, Corbani S, Haidar Z, et al. The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population. *BMC Med Genomics.* 2019;12(1):11. <http://doi.org/10.1186/s12920-019-0474-y>
16. Salfati EL, Spencer EG, Topol SE, et al. Re-analysis of whole-exome sequencing data uncovers novel diagnostic variants and improves molecular diagnostic yields for sudden death and idiopathic diseases. *Genome Med.* 2019;11(1):83. <http://doi.org/10.1186/s13073-019-0702-2>
17. Tan NB, Stapleton R, Stark Z, et al. Evaluating systematic reanalysis of clinical genomic data in rare disease from single center experience and literature review. *Mol Genet Genomic Med.* 2020;8(11):e1508. <http://doi.org/10.1002/mgg3.1508>
18. 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(1):33. <http://doi.org/10.1038/s41525-020-00140-1>
19. Matalonga L, Hernández-Ferrer C, Piscia D, et al. Solving patients with rare diseases through programmatic reanalysis of genome-phenome data. *Eur J Hum Genet.* 2021;29(9):1337-1347. <http://doi.org/10.1038/s41431-021-00852-7>
20. Ji J, Leung ML, Baker S, Deignan JL, Santani A. Clinical exome reanalysis: current practice and beyond. *Mol Diagn Ther.* 2021;25(5):529-536. <http://doi.org/10.1007/s40291-021-00541-7>
21. Wright CF, McRae JF, Clayton S, et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet Med.* 2018;20(10):1216-1223. <http://doi.org/10.1038/gim.2017.246>
22. Deignan JL, Chung WK, Kearney HM, et al. Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2019;21(6):1267-1270. <http://doi.org/10.1038/s41436-019-0478-1>
23. Zurek B, Ellwanger K, Vissers LELM, et al. Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. *Eur J Hum Genet.* 2021;29(9):1325-1331. <http://doi.org/10.1038/s41431-021-00859-0>
24. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46(D1):D1062-D1067. <http://doi.org/10.1093/nar/gkx1153>
25. ClinVar submissions – NCBI – NIH. Accessed February 9, 2022. <https://www.ncbi.nlm.nih.gov/clinvar/submitters>
26. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med.* 2012;367(20):1921-1929. <http://doi.org/10.1056/NEJMoal206524>
27. Thevenon J, Duffourd Y, Masurel-Paulet A, et al. Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin Genet.* 2016;89(6):700-707. <http://doi.org/10.1111/cge.12732>
28. Laurie S, Fernandez-Callejo M, Marco-Sola S, et al. From wet-lab to variations: concordance and speed of bioinformatics pipelines for whole genome and whole exome sequencing. *Hum Mutat.* 2016;37(12):1263-1271. <http://doi.org/10.1002/humu.23114>
29. 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(5):405-424. <http://doi.org/10.1038/gim.2015.30>

30. Cogné B, Ehresmann S, Beauregard-Lacroix E, et al. Missense variants in the histone acetyltransferase complex component Gene TRRAP Cause autism and syndromic intellectual disability. *Am J Hum Genet.* 2019;104(3):530-541. <http://doi.org/10.1016/j.ajhg.2019.01.010>
31. Senaratne TN, Quintero-Rivera F. NFIA-related disorder. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. *GeneReviews [Internet]*. University of Washington, Seattle; 1993-2023. Accessed September 06, 2021 <http://www.ncbi.nlm.nih.gov/books/NBK542336/>
32. Koenig M, Tranchant C, Anheim M. Autosomal recessive cerebellar ataxia 3 due to homozygote c.132dupA mutation within the ANO10 gene—reply. *JAMA Neurol.* 2015;72(2):239-240. Published correction appears in *JAMA Neurol.* 2015;72(4):481. <https://doi.org/10.1001/jamaneurol.2014.3921>
33. Hamdan FF, Daoud H, Piton A, et al. De novo SYNGAP1 mutations in nonsyndromic intellectual disability and autism. *Biol Psychiatry.* 2011;69(9):898-901. <http://doi.org/10.1016/j.biopsych.2010.11.015>
34. de Boer E, Yaldiz B, Denommé-Pichon AS, et al. Genome-wide variant calling in reanalysis of exome sequencing data uncovered a pathogenic TUBB3 variant. *Eur J Med Genet.* 2022;65(1):104402. <http://doi.org/10.1016/j.ejmg.2021.104402>
35. Breuss M, Heng JI, Poirier K, et al. Mutations in the  $\beta$ -tubulin gene TUBB5 cause microcephaly with structural brain abnormalities. *Cell Rep.* 2012;2(6):1554-1562. <http://doi.org/10.1016/j.celrep.2012.11.017>
36. Giunta C, Rohrbach M, Fauth C, Baumann M. FKBP14 kyphoscoliotic Ehlers-Danlos syndrome. In: Adam MP, Everman DB, Mirzaa GM, et al., eds. *GeneReviews [Internet]*. University of Washington, Seattle; 1993-2023. Accessed September 08, 2021 [online]. <http://www.ncbi.nlm.nih.gov/books/NBK541503/>. Seattle, (WA): University of Washington, Seattle.
37. Wright CF, Eberhardt RY, Constantinou P, et al. Evaluating variants classified as pathogenic in ClinVar in the DDD Study. *Genet Med.* 2021;23(3):571-575. <http://doi.org/10.1038/s41436-020-01021-9>
38. Li H. Improving SNP discovery by base alignment quality. *Bioinformatics.* 2011;27(8):1157-1158. <http://doi.org/10.1093/bioinformatics/btr076>
39. Tian S, Yan H, Kalmbach M, Slager SL. Impact of post-alignment processing in variant discovery from whole exome data. *BMC Bioinformatics.* 2016;17(1):403. <http://doi.org/10.1186/s12859-016-1279-z>
40. Yaay K, Lecoquierre F, Baert-Desurmont S, et al. 'Genome Alert!': a standardized procedure for genomic variant reinterpretation and automated genotype-phenotype reassessment in clinical routine. *Genet Med.* 2022;24(6):1316-1327. <http://doi.org/10.1016/j.gim.2022.02.008>