P11.052.B - Getting the most out of Exome Sequencing data

J. López[1], M.J. Roca[1], L. González[1], D. Trujillano[1], C. Camprubí[1], I. Royo[1], X. Pinasa[1], L. Aparicio[1], R. Nava[1], A. Torrents[1]
All authors contributed equally to this study
[1] Reference Laboratory Genetics, L'Hospitalet de LLobregat, Spain.

REFERENCE LABORATORY GENETICS

INTRODUCTION

Whole Exome Sequencing (WES) is currently a solid diagnostic tool for heterogeneous genetic diseases. However, a partial and insufficient analysis of WES data could lead to misdiagnosis. We compiled 13 cases, most of them with previous negative WES results, in which a thorough reanalysis of the data yielded new candidate genes with diagnostic potential.

MATERIALS AND METHODS

Our cohort included 300 patients with syndromic neurodevelopmental abnormalities. WES analyses were guided by HPO terms and custom gene panels based on literature. For those inconclusive results, a secondary and more objective inspection was performed.

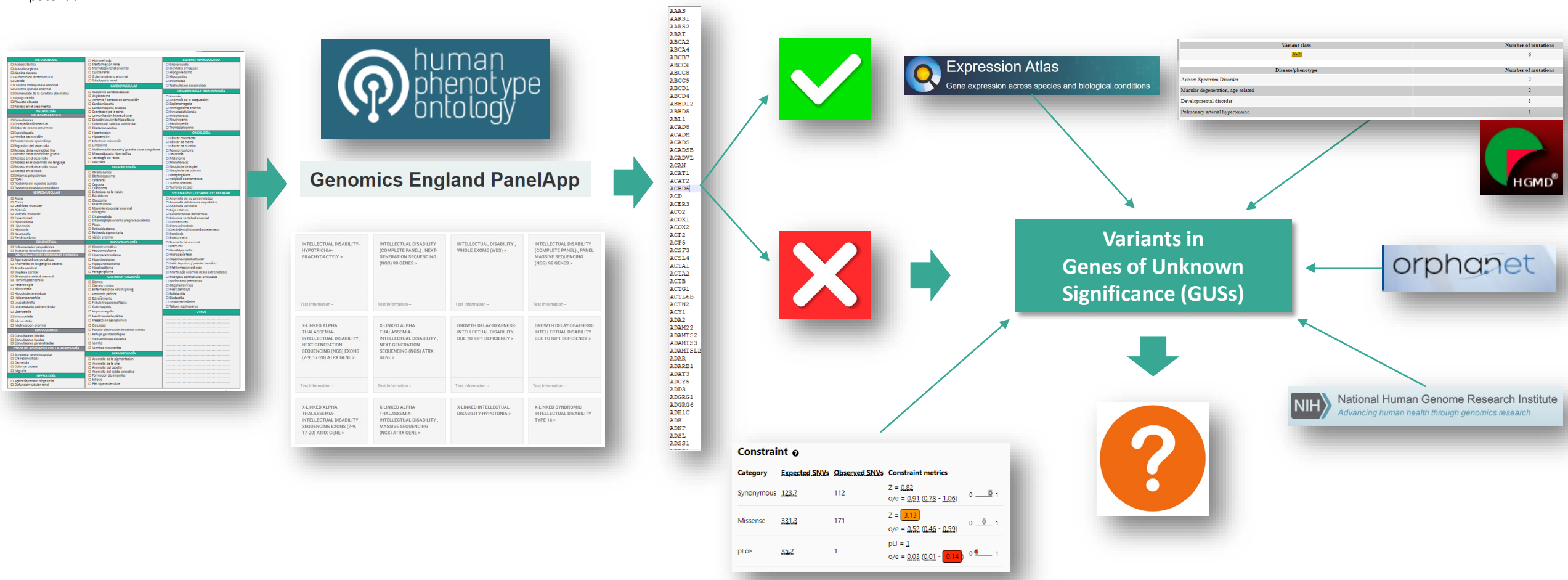


Figure 1. From clinical information supported by the geneticists, we transform the information to HPO terms, to PanelApp or curated NGS panels. When conclusive results were not obtained (red X) variants present in GUSs were analysed. For these variants, we consider information from different databases, such as expression data, intolerance to Loss of Function or missense changes, and clinical databases beyond OMIM. Segregation data of variants in these genes are essential to know the likely clinical association (orange question mark).

RESULTS: Atypical phenotype and new phenotypic associations

Case	Gender	Age	Clinical information of the patients	Gene of interest	Variant	Segregation	Findings
I	F	8	Mild intellectual disability, global developmental delay, autism, NO seizures, **incisor macrodontia, synophrys, bulbous nose**	*SCN2A*	c.3631G>A;p.E1211K (Ht)	*De novo*	Atypical phenotype
II	M	0,5	**Hemi-microtia, hemi-dilatation of the renal pelvis, hemi-micrognathia, dysplastic corpus callosum**	*DPYS*	c.1078T>C;p.W360R (Hm)	NA	New phenotypic association
IIIa	F	Prenatal	**Echogenic fetal bowel, edema, hypoplastic left heart, ventriculomegaly, increased nuchal translucency, cranial asymmetry, aplasia/hypoplasia of the nasal bone, polyhydramnios**	*MVK*	c.1006G>Ap.G336S (Hm)	Parental inheritance	New phenotypic association
IIIb	F	Newborn	Microcephaly, **patent ductus arteriosus, ventriculomegaly, intracerebral periventricular calcifications,** hepatosplenomegaly				

Table 1. Conclusive cases with pathogenic variants in genes which do not explain completely the phenotype of the patients. **In bold** are highlighted the phenotypic findings without previous relationship for these genes in the literature consulted. **Case I** highlights mild dysmorphic features and absence of seizures caused by a *de novo* pathogenic variant in *SCN2A* gene (rare but previously described for this gene in PMID: 24579881) (figure 2). **Case II** presents a very unusual phenotype for *DPYS* gene (figure 3). Note that only the left side of the body was affected in the patient. **Case III (patients IIIa and IIIb)** presents very severe phenotypes resulting in ante- and perinatal death (figure 4). We highlight the echographic prenatal findings related to *MVK* gene. F (Female). M (Male). Hm (Homozygous state). Ht (Heterozygous state). NA (Not available).

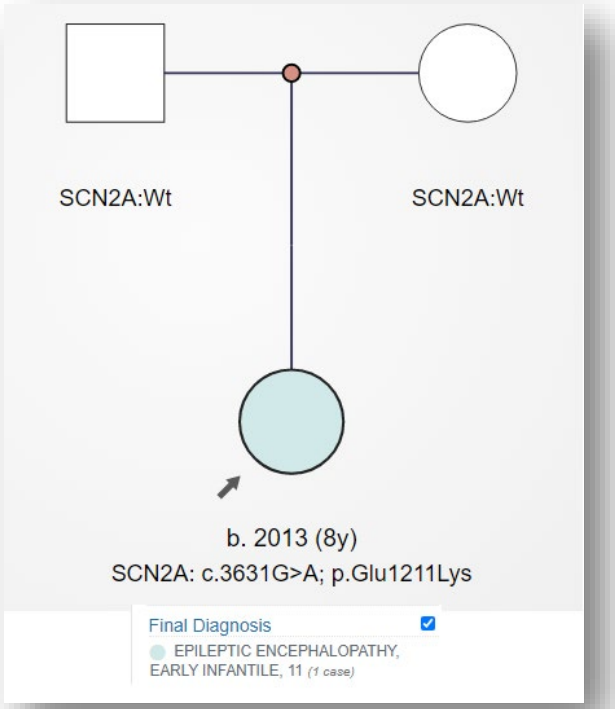


Figure 2. Familial segregation in case I. Wt (Wild type). Age of the proband was added (b=birth date).

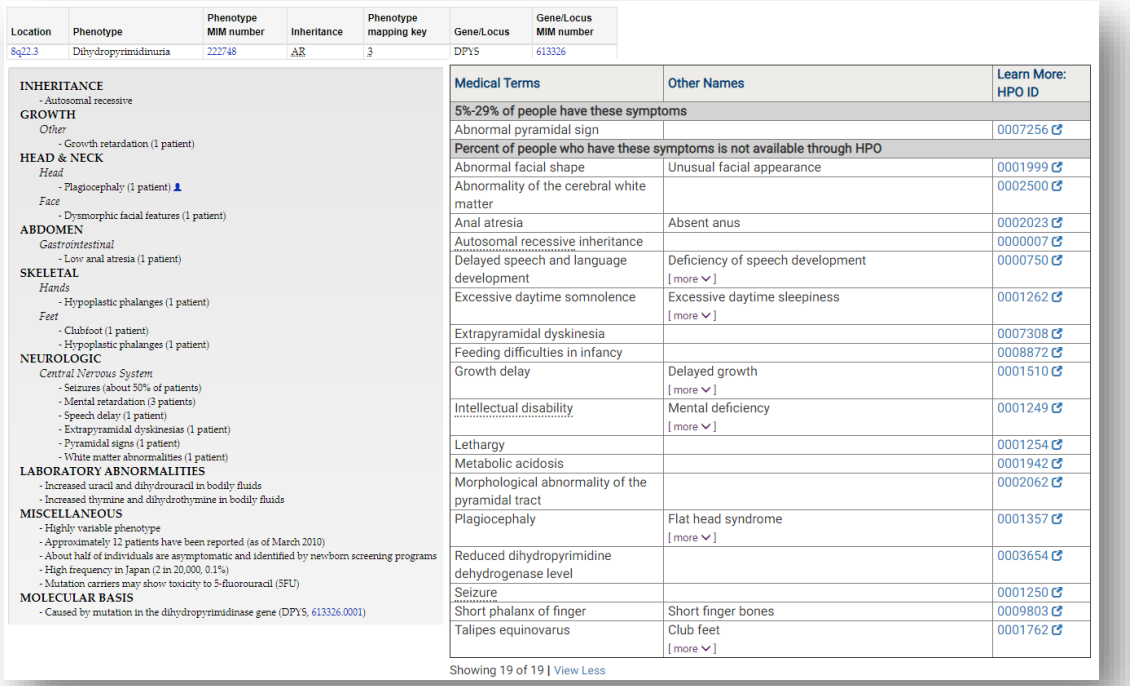


Figure 3. Typical phenotypic findings related to DPYD deficiency (Dihydropyrimidinuria), from OMIM and GARD databases.

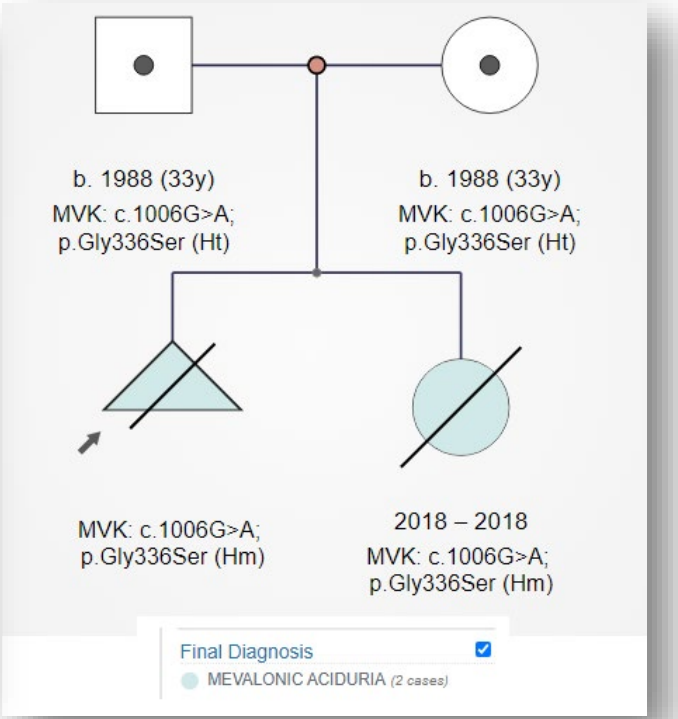
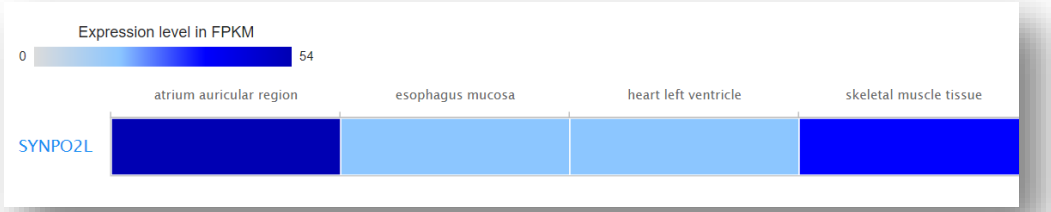


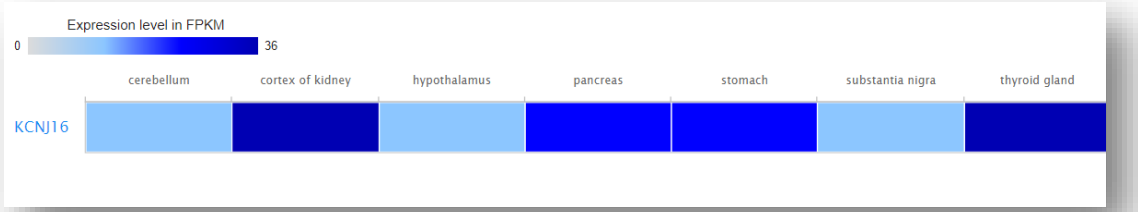
Figure 4. Familial segregation of cases IIIa and IIIb. Hm (Homozygous), Ht (Heterozygous).

RESULTS: GUSs

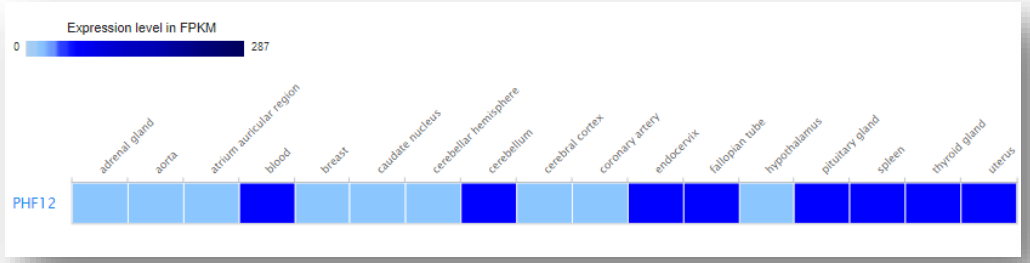
Case	Gender	Age	Clinical information of the patients	Gene of interest	Variant	Segregation/ Disposition	Protein coding	Expression	LoF intolerance: pLI/pRec/pNull values[+]	Previous knowledge	Genes' references
IV	F	6	**Hypokalemia, polydipsia, muscle weakness, abnormal EKG, polyuria, Bartter syndrome suspicion**	*SYNPO2L*	c.631C>T; p.Q211*	NA	Synaptopodin 2-like	Heart and muscle	0/0,9/0,1: High (autosomal recessive predicted)	Likely association with Abnormal EKG, atrial fibrillation	PMID: 29748316, 22544366, 29290336, 28416818, 30061737, 29892015
				KCNJ16	c.631C>T; p.R211*	Cis	potassium inwardly rectifying channel subfamily J member 16	Cortex of kidney and others	0/0/0,98: Likely LoF tolerance	Likely association with cardiopathy and Thyrotoxic periodic paralysis	PMID: 30821013, 30662450, 22863731
					c.779A>T; p.E260V	Cis					
					c.509G>C; p.G170A	trans					
V	M	7	Autism, borderline intellectual disability, **abnormal skull morphology, prominent glabella, high anterior hairline, hypertelorism, epicanthus, smooth philtrum, thin upper lip vermilion, overfolded helix, camptodactyly of toe, abnormality of the cervical spine, inguinal hernia, cryptorchidism, hypermetropia, astigmatism, optic disc pallor.**	*PHF12*	c.2059C>T; p.Q687*	*De novo*	PHD finger protein 12	Cerebellum, cortex, blood, others	1/0/0: High (autosomal dominant predicted)	Likely association with autism and developmental disorder	PMID: 31981491,28135719, 28191890
VI	F	5	Neurodevelopmental delay, delayed speech and language development, autism, **motor delay, intrauterine growth retardation, stereotypy, polyphagia, thin upper lip vermilion, smooth philtrum.**	*XKR6*	c.1672_1675del; p.T558Afs*7	NA	Kell blood group complex subunit-related family	Cerebellum	0,9/0/0: High (autosomal dominant predicted)	Likely association with autism	PMID: 28191890, 25363768

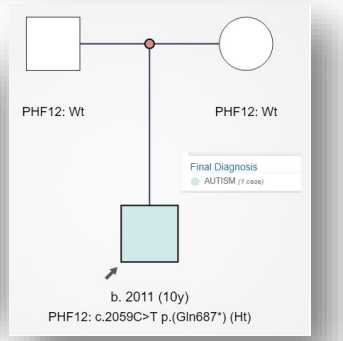
Table 2. Cases with variants in Genes of Unknown Significance (GUSs) that could explain the phenotype. **In bold** are highlighted the phenotypic findings without previous relationship for these genes in the literature consulted. **Case IV,** has two candidate genes to explain the clinical findings of the patient according to their expression and previous literature knowledge (figure 5-6). **Case V** highlights the cerebellum expression of *PHF12* gene, the theorical intolerance to Loss of Function (LoF), the *de novo* origin and the likely association with autism. **Case VI** has similar characteristics that case V but we do not have segregation information. [(+)] Information of Loss of Function (LoF) intolerance metrics from gnomAD database: PMID: 27535533.





Figures 5-6. Tissue expression for *SYNPO2L* and *KCNJ16* genes. *KCNJ16* shows a high expression in tissues that could be related to the water-electrolyte imbalance, and *SYNPO2L* could be related to the patient's muscle weakness.





Figures 7-8. Tissue expression for *PHF12* gene and the familial segregation. Another PHD finger proteins such as *PHF6* or *PHF8* have been associated with similar phenotypes (OMIM: 301900, 300263)

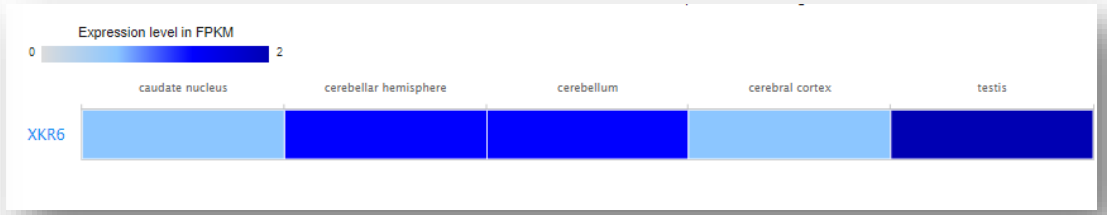


Figure 9. *XKR6* gene shows a relative high expression in cerebellar tissue. However, there is very little information about this family of genes. A precursor of kell blood group (*XK* gene) has been associated with McLeod syndrome, with some neurological implications (OMIM: 300842). Moreover, other members of XKR family (*XKR4* and *XKR9* genes) have been related to autism or schizophrenia (PMID: 28191890, PMID: 24463507).

RESULTS: New pathogenic/likely pathogenic variants

Case	Gender	Age	Clinical information of the patients	Gene of interest	Variant	Segregation	Allele frequency (GnomAD)	HGMD/ClinVar descriptions	Effect	Mode of inheritance	ACMG and ACGS criteria
VII	F	31	Epidermolysis bullosa simplex	*KRT14*	c.922_927+14delinsCTCGTTTGAGGA; p.T308_T310delinsLV	Paternal Inheritance	0	ND	LoF (splicing disrupted)	AD	PVS1+PM2
VIII	F	10	Intrauterine growth retardation, **microcephaly**, **brachycephaly**, low-set ears, microtia, high palate, abnormality of the periorbital region, stellate iris, hypermetropia, astigmatism, perimembranous ventricular septal defect, intellectual disability borderline, specific learning disability, language impairment	*NOTCH3*	c.2968G>T; p.E990*	NA	0	ND	LoF (nonsense)	AD	PVS1+PM2
IX	F	5	triangular face, broad forehead, abnormality of the scalp hair, narrow mouth, short philtrum, abnormal distal phalanx morphology of finger, umbilical hernia, global developmental delay, intellectual disability	*SMARCA2*	c.3638G>C; p.R1213P	NA	0	ND/ 495133 (VOUS)	GoF? (PMID:22822383)	AD	PM2, **PM5**, PP3, **PM1, PP2**
X	F	8	autism, encephalopathy, seizures, intellectual disability, narrow palate, misalignment of teeth, single transverse palmar crease, obesity, language impairment.	*MECP2*	c.890_893delinsCCC; p.Q297Pfs*24	NA	0	ND	LoF (frameshift)	XL	PVS1+PM2
XI	F	5	global developmental delay, delayed speech and language development, delayed social development, abnormal facial shape	*SHANK3*	c.2669_2670dupGC; p.S891Afs*3	*De novo*	0	ND	LoF (frameshift)	AD	PVS1+PM2

Table 3. Cases with pathogenic or likely pathogenic variants not previously described. **In bold** are highlighted the phenotypic findings without previous relationship for these genes in the literature consulted. In **case VII,** the pathogenic variant was inherited from the affected father (see segregation in figure 10). This change disrupts the canonical splice site from intron 4 of *KRT14* gene. In **case VIII,** we detected a pathogenic variant in *NOTCH3* gene, confirming a diagnosis of atypical lateral meningocele syndrome in the patient, with microcephaly and brachycephaly. In **case IX,** we identified a likely pathogenic missense variant in a conserved domain (PM1) of *SMARCA2.* This gene has a high intolerance to missense changes (PP2), so, the pathological mechanism could be associated with a Gain of Function (GoF) (see figure 11). In **case XI,** we identified a new frameshift mutation in *SHANK3* gene, with a *de novo* origin (see Figure 12). Classification of the variants based on ACMG and ACGS criteria (Ellard S. et al., 2020). ND (Not described). NA (Not available).

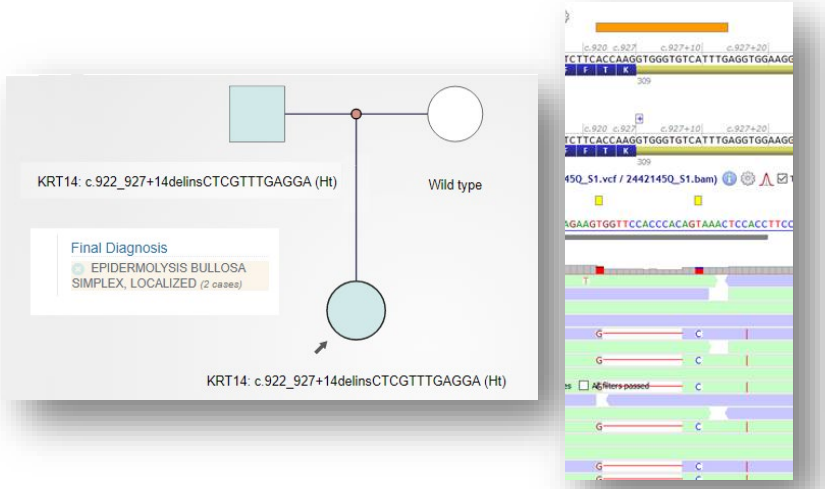


Figure 10. Left: segregation of the variant in *KRT14* gene. Right: visualisation of the indel variant, breaking the canonical splice site.

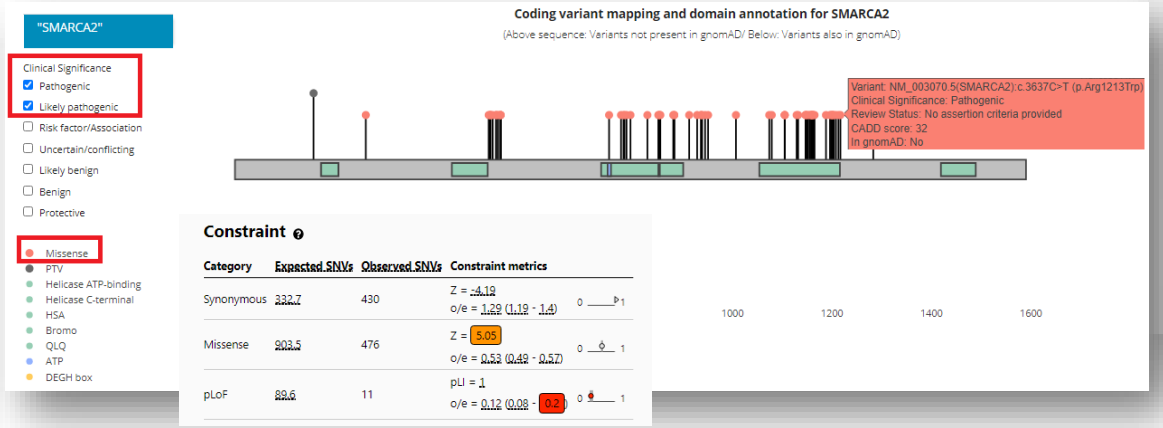


Figure 11. The majority of pathogenic variants in *SMARCA2* gene are missense and located in the central region of the gene (ClinVar database), where the variant of the patient is located**.** Z-score metric from gnomAD (below) shows a high value (5,05), compatible with a missense variant intolerance.

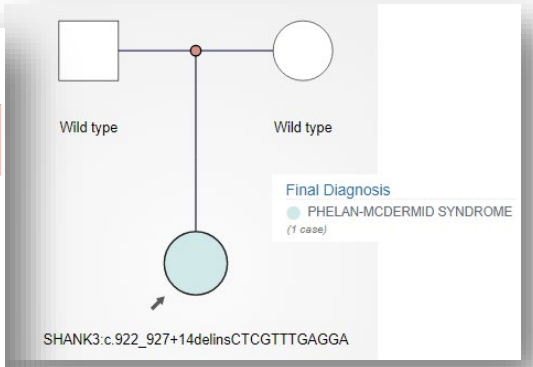


Figure 12. Segregation of variant in *SHANK3* gene.

RESULTS: Two pathologies in a family

Case	Gender	Age	Clinical information	Gene of interest	Variant	Segregation
XII	F	Newborn	Enlarged kidney, renal dysplasia, renal cyst, hyperechogenic kidneys, oligohydramnios, neonatal hypotonia	*BBS1*	c.1169T>G;p.M390R	Hm
				ALG8	c.96-1G>C	Paternal inheritance
XIII	M	14	Intellectual disability, microcephaly	*POGZ*	c.1535delA; p.H512Lfs*2	*De novo*
				TRIO	c.2755-2A>G	Maternal inheritance

Table 4. Both patients have been identified with two pathogenic variants in different genes that could explain a particular mixed phenotype. In **case XII,** we detected a familial pathogenic variant in *ALG8* gene that explains the renal cysts present in some relatives (see figure 13) likely with an adult onset. Moreover, the proband had a homozygous pathogenic variant in *BBS1* gene producing a more severe phenotype. To date, no physical interactions or digenic inheritance between these genes (or their proteins) have been reported (figure 14). In **Case XIII,** we present a familial case of mild intellectual disability and microcephaly. In addition to the familial pathogenic variant in *TRIO* gene, the proband also had a *de novo* pathogenic change in *POGZ,* resulting in a more severe phenotype (Figure 15).

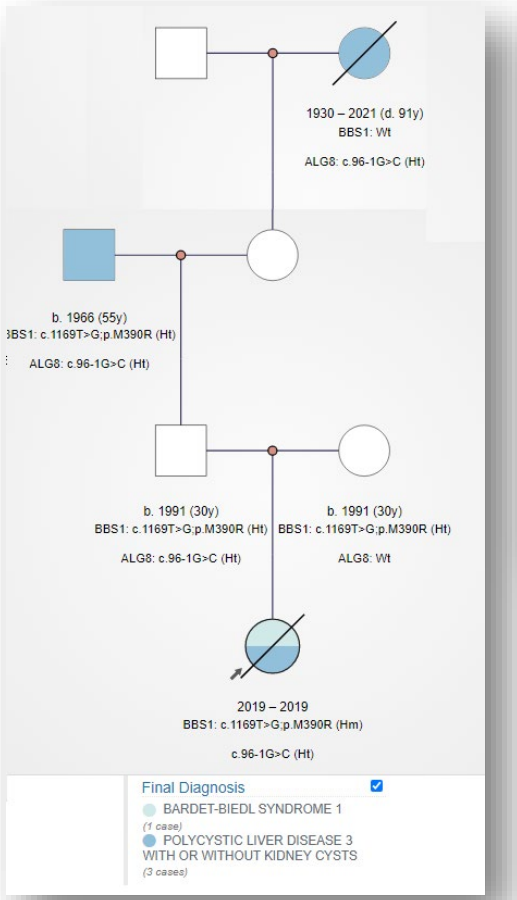


Figure 13. Segregation of the variants in case XII, with some affected relatives.

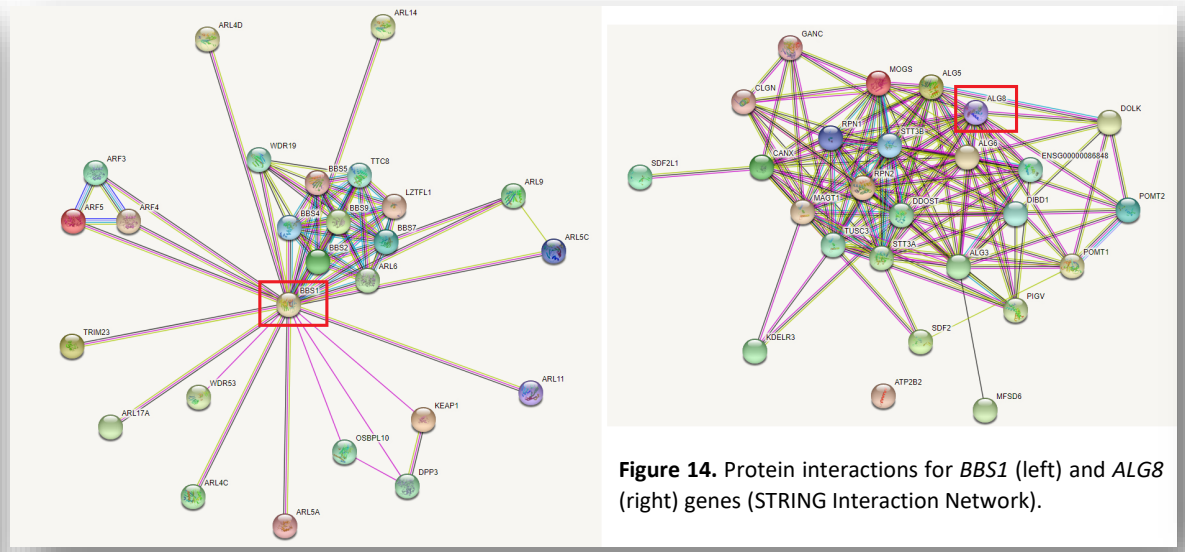


Figure 14. Protein interactions for *BBS1* (left) and *ALG8* (right) genes (STRING Interaction Network).

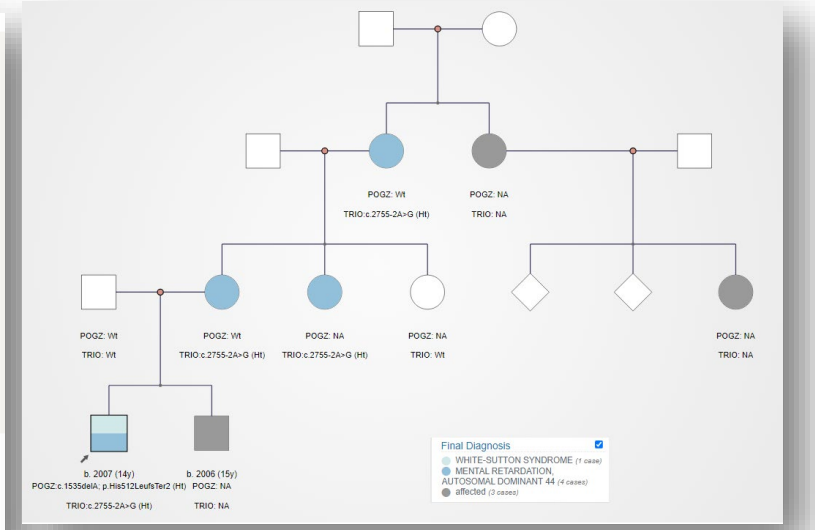


Figure 15. Familial segregation for *POGZ* and *TRIO* genes. In grey are marked affected relatives without molecular confirmation.

CONCLUSIONS

The results highlight the clinical diagnostics potential of a thorough **reanalysis of previously negative WES cases**. Based on the identification in 5% of cases of **new candidate genes, the increase of phenotypic spectrum of the diagnosed disease** (even reporting cases with **a dual diagnosis**) and the findings **of novel pathogenic variants**, we recommend **going beyond the basic guided genetic analysis for HPO terms.** Which is particularly relevant in cases with a strong suspicion of an underlying genetic disease and a previous inconclusive WES result.

This communication aims to remark the **high complexity of the genetic studies** in an elevated percentage of cases, where a unique mendelian cause in a well-known change or gene cannot explain the pathology of the patients.

Unconclusive results need a detailed analysis. The use of research data is necessary to **select new candidate genes,** and the analyst should keep in mind the possibility of **atypical phenotypes**, evaluating the biological risk of pathogenicity of the variants in addition to the clinical information available for each gene. Finally, an **accurate classification of variants** is essential to select the best candidate variants.