Biallelic loss of function variants in *WBP4*, encoding a spliceosome protein, result in a variable neurodevelopmental syndrome

Eden Engal[1,2*], Kaisa Teele Oja[3,4*], Reza Maroofian[5], Ophir Geminder[1,2], Thuy-Linh Le[6], Pauline Marzin[7], Anne Guimier[6,7], Evyatar Mor[8], Naama Tzvi[9], Naama Elefant[9], Maha S. Zaki[10], Joseph G. Gleeson[11,12], Kai Muru[3,4], Sander Pajusalu[3,4], Monica H. Wojcik[13], Divya Pachat[14], Marwa Abd Elmaksoud[15], Won Chan Jeong[16], Hane Lee[16], Peter Bauer[17], Giovanni Zifarelli[17], Henry Houlden[5], Muhannad Daana[18], Orly Elpeleg[9,19], Jeanne Amiel[6,7], Stanislas Lyonnet[6,7], Christopher T. Gordon[6], Tamar Harel[9,19], Katrin Õunap[3,4#], Maayan Salton[1#], Hagar Mor-Shaked[9,19#]

[1] Department of Biochemistry and Molecular Biology, The Institute for Medical Research Israel-Canada, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem 9112102, Israel

[2] Department of Military Medicine and "Tzameret", Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

[3] Genetics and Personalized Medicine Clinic, Tartu University Hospital, Tartu, Estonia

[4] Institute of Clinical Medicine, University of Tartu, Tartu, Estonia

[5] Department of Neuromuscular Disease, UCL Queen Square Institute of Neurology, London, UK

[6] Laboratory of embryology and genetics of human malformations, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1163, Institut Imagine and Université Paris Cité, Paris 75015, France

[7] Service de Médecine Génomique des Maladies Rares, Hôpital Necker-Enfants Malades, AP-HP, Paris 75015, France

[8] Department of computer science, Ben-Gurion University of the Negev

[9] Department of Genetics, Hadassah Medical Organization, Jerusalem, Israel

[10] Department of Clinical Genetics, Human Genetics and Genome Research Institute, Cairo, Egypt

[11] Department of Neurosciences, University of California, San Diego, La Jolla, USA

[12] Rady Children's Institute for Genomic Medicine, San Diego, La Jolla, USA

[13] Broad Institute of MIT and Harvard, Cambridge, MA

[14] Department of Medical Genetics, Aster MIMS (Malabar Institute of Medical Sciences)-Calicut, Kerala, India

[15] *Neurology Unit, Department of Pediatrics, Faculty of Medicine, Alexandria University, Alexandria, Egypt*

[16] *3billion, Seoul, South Korea (or Republic of Korea)*

[17] *CENTOGENE GmbH, Am Strande 7, 18055 Rostock, Germany*

[18] *Child Development Centers, Clalit Health Care Services, Jerusalem, Israel*

[19] *Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel*

[*],[#] These authors contributed equally.

Corresponding authors:

Hagar Mor-Shaked, Ph.D.

Department of Genetics

Hadassah-Hebrew University Medical Center

POB 12000

Jerusalem, Israel

+(972)-2-6776931 (office)

+(972)-2-6777618 (fax)

Email: hagarmor@hadassah.org.il

Keywords: pre-mRNA splicing, spliceosome, syndromic neurodevelopmental disorder, *WBP4*

ABSTRACT

Over two dozen spliceosome proteins are involved in human diseases, also referred to as spliceosomopathies. WBP4 (WW Domain Binding Protein 4) is part of the early spliceosomal complex and has not been previously associated with human pathologies in the Online Mendelian Inheritance in Man (OMIM) database. Through GeneMatcher, we identified ten individuals from eight families, with a severe neurodevelopmental syndrome featuring variable manifestations. Clinical manifestations included hypotonia, global developmental delay, severe intellectual disability, brain

abnormalities, musculoskeletal and gastrointestinal abnormalities. Genetic analysis revealed five different homozygous loss-of-function variants in *WBP4*. Immunoblotting on fibroblasts from two affected individuals with different genetic variants demonstrated a complete loss of protein, and RNA sequencing analysis uncovered shared abnormal splicing patterns, including in genes associated with abnormalities of the nervous system, potentially underlying the phenotypes of the probands. We conclude that biallelic variants in *WBP4* cause a developmental disorder with variable presentations, adding to the growing list of human spliceosomopathies.

MAIN TEXT

The spliceosome is a complex of RNA and proteins responsible for promoting accurate splicing. More than two dozen spliceosome proteins are known to be involved in human diseases, also referred as spliceosomopathies. Despite the fundamental function of the spliceosome in all tissues, spliceosomopathies often involve defects restricted to specific tissues.[1] For example, mutations in SNRNP200 [MIM: 601664], a widely expressed component of the U4/U6 complex, are a known cause of Retinitis pigmentosa 33 [MIM: 610359], which exclusively affects the retina.[2] Another example is mutation of *SNRPB* [MIM: 182282] causing cerebrocostomandibular syndrome [MIM: 117650].[3] These observations suggest that particular components of the spliceosome may play a crucial role in accurately splicing specific targets, which can lead to the development of distinct pathologies, and/or that certain cell types are more sensitive than others to perturbations of ubiquitous spliceosomal functions.

WBP4 (WW Domain Binding Protein 4), previously known as FBP21, is a 376-amino acid spliceosome protein that contains a zinc finger motif and two tandem WW domains. It was first detected as part of the early spliceosomal complex, where it interacts with the U2-associated protein SF3B4, SIPP1, and the core splicing protein SmB/B'. These interactions are facilitated by the transient multivalent recognition of proline-rich sequences by the WW domains.[4–7] Later it was found to be present in the Spliceosome B complex and also to bind and inhibit the RNA helicase SNRNP200.[8,9] In the transition

from the B to the Bact complex, SNRNP200 unwinds the U4/U6 duplex, releasing the U4 snRNP, and enabling the U6 snRNA to base-pair with the U2 snRNA, forming a catalytically important stem–loop.[10,11] Thus, inhibition of SNRNP200 plays a crucial role in splice site selection, which positions WBP4 as a key player in determining the outcome of splicing. Additionally, WBP4 was shown to enhance splicing in vitro and in vivo,[6] and to regulate alternative splicing,[12] but has not been previously associated with human pathologies on OMIM.

Through GeneMatcher[13], a total of ten affected individuals from eight families were identified to carry homozygous LOF variants in *WBP4* [MIM: 604981] (Table 1 and Figure 1A, 1C). All families had given informed consent, and the study was approved by institutional review boards of the corresponding institutions (details in supplemental methods). The most common clinical features among the affected individuals in the eight families were muscular hypotonia, global developmental delay, dysmorphic facial features, and feeding difficulties or failure to thrive (Table 1). Moderate to severe intellectual disability was present in all affected individuals who were old enough for evaluation. Aplasia or hypoplasia of corpus callosum and atypical behavior were present in almost half of them. In some cases, congenital abnormalities of the heart, gastrointestinal, skeletal, and/or urogenital system were evident. The clinical features of the cohort are described in detail in Table S1, and summaries of affected individuals (where available) are provided below.

In Family 1, the proband had a severe intellectual disability and autism. He was a male child born to a consanguineous Arab Muslim family (Figure 1A, F1-III:1) with another affected offspring (individual F1-III:4; fetus with atrioventricular (AV)-canal, suspected absence of thymus, and intrauterine growth restriction). Quad exome sequencing was performed, but no pathogenic variants in genes known to be associated with neurological disorders were found, or any other known pathogenic events in both siblings. As the consanguineous background was suggestive of autosomal recessive inheritance, we focused on rare homozygous variants. A potentially harmful homozygous frameshift variant in *WBP4* was found in both siblings (NM_007187.5:c.499del, p.(Thr167Profs*4)), predicted to result in a premature stop codon. Segregation study by Sanger sequencing revealed that the parents and two

unaffected siblings were heterozygous carriers, and the affected individuals were homozygous for the variant (Figure 1A and Figure S1).

The c.499del variant was found in three additional families (Families 3-5) of Egyptian origin and was present in six affected individuals in total. Haplotype analysis suggested a common haplotype, containing *WBP4*, of approximately 1.5Mb shared between four affected individuals in Families 1, 3, and 4 (Figure S2). More cases and controls are needed to confirm whether it is a distinct haplotype of this specific frameshift variant or a common uninformative haplotype.

The proband from Family 2 (F2-II:1) had a severe phenotype. He was a male born as the first child to non-consanguineous parents of European descent. Bilateral hydronephrosis, intrauterine growth retardation, and mild oligohydramnios were seen during pregnancy on ultrasound. The child was born at 35+5 weeks of gestation with growth parameters below the 3rd percentile (Table S1). He had multiple congenital anomalies – anal atresia with fistula into the bladder, distal hypospadia and cryptorchidism, bilateral hydronephrosis, congenital heart defect - atrial septal defect, agenesis of corpus callosum (CCA) and dysmorphic features (dolichocephaly, proptosis, hypertelorism, depressed nasal bridge, microretrognathia; Figure 1B). Clinically, the VACTERL association was diagnosed. He presented failure to thrive, bilateral neurosensory hearing loss, optic nerve atrophy, convergent strabismus, global developmental delay, hypotonia, dysphagia, and frequent infections. Brain MRI confirmed the CCA and frontotemporal widening of subarachnoidal space. Trio exome analysis did not yield a molecular diagnosis and was followed by trio genome sequencing, RNA sequencing from fibroblast cell cultures, and untargeted metabolomics analysis from serum. RNA sequencing revealed a very low expression of *WBP4* (Figure S3), and reanalysis of genome data indicated a homozygous deletion in this region (Figure S4). Homozygous deletion of the three last exons of *WBP4* was validated by Sanger sequencing. The deletion is located in a ~6.2 Mb size region of homozygosity (precise coordinates for deletion chr13:41,074,140-41,090,168 (hg38)).

In family F7, the female proband was born to consanguineous Moroccan parents. Prenatal scans identified growth delay, ventriculomegaly, CCA and craniosynostosis. In the neonatal period,

craniofacial anomalies were noted (including hypertelorism, cloverleaf skull, and exophthalmos; Figure 1B) along with Hirschsprung's disease and hypotonia. Postnatal cerebral imaging confirmed the CCA and revealed cortical gyration anomalies. She developed epilepsy at 2 years and global developmental delay was evident at 3 years. Singleton exome sequencing of the proband revealed a homozygous frameshift variant in *WBP4*: c.944del; p.(Pro315Glnfs*55). Although this variant falls in the final exon, and therefore potentially leads to escape from nonsense-mediated decay, any truncated protein produced would lack a C-terminal region known to interact with SNRNP200[8].

In the female proband of family F8, several malformations were identified in the neonatal period, including clubfeet, cleft lip and palate, cortical gyration anomalies, and congenital heart defects (ventricular septal defects and coarctation of the aorta), the latter leading to her death in early infancy. Singleton exome sequencing identified a homozygous canonical splice site variant in *WBP4* in the proband: c.440-1G>A (falling in the acceptor site of intron 5).

Overall, 5 different *WBP4* LOF variants were detected across the 8 families (Table 1), all with absent or extremely low frequency (<0.01%) in gnomAD Exomes v2.1.1 and gnomAD genomes v3.1.2.[14] Although the probability of loss of function intolerance (pLI) in gnomAD v2.1.1 is zero for *WBP4*, indicating a high tolerance for heterozygous LOF variants, no homozygous LOF variant is present in this gene in gnomAD v2.1.1 or v3.1.2.[14]

Considering the fact that *WBP4* codes for a spliceosome protein, we wished to understand how splicing is altered in cells lacking WBP4. For this, we took a skin biopsy from two affected individuals F1-III:1 and F2-II:1 and their parents. Primary fibroblasts were grown, and protein was extracted to evaluate the amount of WBP4 in the probands relative to their heterozygous parents and WT primary fibroblasts, CCD-1092Sk. The heterozygous parents showed 25-50% of WBP4 amount relative to the WT fibroblasts and probands with homozygous *WBP4* variants showed total ablation of WBP4 (Figure 2A and B). In addition, we extracted RNA from fibroblasts of the parents as well as probands and conducted RNA sequencing (RNA-seq). *WBP4* mRNA amount was reduced in proband F1-III:1, as is expected by an early termination codon leading to nonsense mediated decay (Figure S5). This process

is acknowledged to exhibit variable efficiency, thus leading to the presence of low levels of *WBP4* mRNA rather than complete ablation.[15] In proband F2-II:1 expression was low, and the transcript terminates at exon 7 out of 10 (NM_007187.5) as expected by the deletion location relative to *WBP4* (Figure S5). Following analysis of the RNA-seq results, we detected 579 genes with changes in gene expression in proband F1-III:1 relative to his parents (FDR<0.05, DEseq2[16]). To understand if these genes contribute to the proband's phenotype, we used the GeneAnalytics tool[17] to identify diseases related to the gene set. The human phenotype ontology pointed at enrichment for skin related phenotypes (Figure S6A). Analyzing the results of the F2-II:1 proband we found 1,068 genes with altered gene expression as compared to his parents (FDR<0.05, DEseq2[16]) with a human phenotype enrichment of skin conditions as well as abnormalities in cardiovascular physiology (Figure S6A). We hypothesize that enrichment for skin diseases, which are not part of the probands' phenotypes, could suggest a difference in gene expression of skin-related genes between the young probands and their adult parents' fibroblasts. Comparing the change in gene expression between the two probands we found 89 overlapping genes (P value<0.00001, Fisher exact test, Figure S6B) with 'abnormality of the face' as the only clinically relevant gene set enrichment (Figure S7, Table S2).

Next, we used rMATS (version 4.1.2) to identify abnormal splicing in the probands.[18] Our analysis identified 5,675 differential splicing events in 2,859 genes in proband F1-III:1 compared to his parents (Figure S8A and Table S3). All types of splicing events were present: skipped exons (2,445 events), alternative 5' splice site (654 events), alternative 3' splice site (629 events), mutually exclusive exons (1,268 events) and intron retention (679 events). To check if the differentially spliced genes are related to the proband's phenotype, we performed gene set enrichment analysis using the above-mentioned tools and found enrichment for abnormalities of the nervous system (Figure S8B). Comparing proband F2-II:1 to his parents we identified 2,868 differential splicing events in 1,555 genes (Figure S8C and Table S3). Again, all types of splicing events were identified: skipped exons (1,361 events), alternative 5' splice site (270 events), alternative 3' splice site (173 events), mutually exclusive exons (478 events) and intron retention (586 events). Gene enrichment analysis pointed to the nervous system, the head

or neck and the skeletal system, similar to proband F1-III:1 (Figure S8D). *WBP4* is expressed in all adult tissues, and during development, including high expression in the fetal brain[19] (Figure S9A and B), consistent with the abnormalities of the nervous system observed in individuals with *WBP4* syndrome (Table S1).

Comparing the abnormal splicing events in the two probands resulted in 619 genes (Figure 3A, Figure S10A-C, and Table S3, P value<0.00001, Fisher exact test), with enrichment for abnormalities of the nervous system and higher mental function, suggesting that the overlapping differentially spliced genes are related to the common phenotypes of the probands (Figure 3B and Table S4). Among the differentially spliced genes related to the probands' phenotype, we identified for example *HUWE1*, *SMARCA2*, *SMARCC2*, *HNRNPH1* and *TCF4* (Table S3 and Table S4). To understand if a connection exists between the regulation of expression to that of splicing by WBP4 we compared the genes changing in expression to that of splicing. We found a small but significant overlap that could be the result of an abnormal splicing event leading to a change in gene expression, such as an inclusion of an exon or a retained intron harboring a stop codon (Figure S10A-C).

Given that the parents of the probands possess less than half of control levels of WBP4 (Figure 2), and recognizing the sensitivity of fibroblasts to age-related changes, we conducted an additional analysis comparing the probands to available RNA-seq data of age-matched healthy fibroblasts.[20] The principal components of the quantile normalized and standardized reads per exon show a distinction from four healthy age-matched controls (Figure S11). Following analysis of the RNA-seq results, we detected 236 differentially expressed genes (Figure S12 Table S2). In addition, our splicing analysis identified 3,564 differential splicing events in 1,836 genes in probands compared to age-matched healthy fibroblasts with enrichment for abnormalities of the nervous system (Figure 3C, D and S13, Table S3 and S4). All types of splicing events were present: skipped exons (746 events), alternative 5’ splice site (227 events), alternative 3’ splice site (150 events), mutually exclusive exons (1968 events), and intron retention (473 events). In order to determine the clinical relevance of aberrantly spliced genes that are shared by the two probands (F1-III:1 and F2-II:1) and were also detected when comparing the

probands to parents and to age-related controls (total of 317 genes, Figure S13D), we performed gene-phenotype relationship analysis. We analyzed the association between the aberrantly spliced gene symbols to 19 phenotypes found in at least three affected individuals (Table 1). The search resulted in 186 (58.6%) directly related aberrantly spliced genes. The top 10 highest ranking hits were – *NFIX*, *TRIO, SMARCC2, PIGT, QRICH1, CDK16, DYNC1H1, PYCR1, HRAS, TCF4* (Table S5), with at least 14 phenotype matches out of 19. This suggests that the abnormally spliced genes are strongly connected to the probands' phenotypes.

Overall, the significant and shared alterations in splicing patterns in both probands, despite the different ethnic backgrounds, point to a similar cause. As suggested by the role of WBP4 in the inhibition of SNRNP200 and its exclusive presence in the B complex, where it acts directly before spliceosomal catalytic activation, depletion of WBP4 in both probands leads to a partially overlapping pattern of abnormal splicing, suggesting a shared pathomechanism of WBP4 loss of function. In addition, monitoring the expression of SNRNP200 in the samples used in this study, we noticed an increase in one of the probands (F2-II:1) but not the other (F1-III:1) (Figure S14). We therefore conclude that the expression level of SNRNP200 does not contribute to the aberrant splicing of the 317 genes identified in Figure S13D.

The spliceosome is comprised of small noncoding RNA molecules (U1, U2, U4, U5 and U6) as well as proteins.[21–24] Among the spliceosome proteins, WBP4 is part of a group that plays a critical role in splice site decisions. These B-complex-specific proteins act directly before spliceosomal catalytic activation, fixing the specific splicing pattern of a substrate through irreversible loss of U4. In the transition from the B to the Bact complex, the RNA helicase SNRNP200 unwinds the U4/U6 duplex, releasing the U4 snRNP and enabling the U6 snRNA to base-pair with the U2 snRNA and forming a catalytically important stem–loop. Therefore, WBP4 is important for spliceosomal activation.

Precise and fine-tuned gene expression is crucial for proper brain development. Neurodevelopment regulatory pathways are coordinated in space and time to produce a well-connected neuronal network.[25] Over two dozen spliceosome proteins have been linked to human diseases, which are

collectively referred to as spliceosomopathies. The tissue-specific nature of several spliceosomopathies suggests that spliceosome components may not be functionally equivalent in all cell types. Studying spliceosomopathies can increase our understanding of the spliceosome and its functions.

Here we report a highly variable neurodevelopmental syndrome caused by homozygous loss of function variants in *WBP4*, in eight different families. The affected individuals had different severity of clinical presentation. Individuals F1-III:1 and F2-II:1 presented the two sides of the spectrum, with F2-II:1 displaying a severe phenotype leading to multiple malformations and premature death and F1-III:1 showing a less severe phenotype, with mainly intellectual disability and motor delay. Nevertheless, the two individuals have shared alterations in splicing patterns, stemming from WBP4 loss of function. These phenotypic differences are unlikely to be due to the nature of the genetic variants (large deletion in F2-II:1 vs. a frameshift variant in F1-III:1), because both probands were shown to have complete loss of WBP4, and other individuals in the cohort with stop-gain or frameshift variants also have a multi-system disorder. Phenotypic variability could also be a characteristic of spliceosomopathies, as previously shown for *CWC27* variants, which lead to a spectrum of conditions including retinal degeneration, short stature, craniofacial abnormalities, brachydactyly, and neurological defects.[26] Another possibility is that proband F1-III:1 carries variants that enable protection against a multi-system syndrome. Of note, no other potentially causal variants were identified through the reanalysis of the whole genome sequencing data of individual F2-II:1.

Seven out of the top ten abnormally spliced genes with the highest phenotype match are known to contain variants implicated in autosomal dominant disease. This might indicate that splicing alterations resulting from biallelic loss of *WBP4* mimic the effect of the cumulative alteration of multiple developmental disorder genes (even if the effect on each of these targets may be individually subtle).

One individual F2-II:1 presented the VACTERL association, which is comprised of at least three of the following characteristic features – vertebral defects, anal atresia, cardiac defects, tracheoesophageal

fistula, renal anomalies, and limb abnormalities. It has been previously shown that heterozygous loss of WBP11 function causes a variety of overlapping congenital malformations, including cardiac, vertebral, tracheo-esophageal, renal, and limb defects.[27] *WBP11*, similarly to *WBP4*, encodes a component of the spliceosome with the ability to regulate pre-RNA splicing. *Wbp11* heterozygous null mice are small and exhibit defects in the axial skeleton, kidneys, and esophagus. Our results support the conclusion of Martin et al 2020[27] that loss-of-function *WBP11* and *WBP4* variants should be considered as a possible cause of VACTERL association.

A larger cohort of affected individuals is necessary to define the entire spectrum of the *WBP4*-related disorder, and further functional studies will lead to a better understanding of the mechanism of pathogenicity.

ACKNOWLEDGEMENTS

The authors wish to thank the families for their participation in this study.

This work is supported by Estonian Research Council grants PRG471 and PSG774, the Agence Nationale de la Recherche “Investissements d’Avenir” program (ANR-10-IAHU-01), MSDAvenir (Devo-Decode project) and AXA (“Tête et Cœur” project). The Broad Institute Center for Mendelian Genomics (UM1HG008900) is funded by the National Human Genome Research Institute with supplemental funding provided by the National Heart, Lung, and Blood Institute under the Trans-Omics for Precision Medicine (TOPMed) program and the National Eye Institute. MHW is supported by K23HD102589.

AUTHOR CONTRIBUTIONS

H.M.-S., K.T.O., M.S., and E.E. designed the study and wrote the paper. T.H. and K.O. supervised the study, and contributed to writing the paper. E.E. performed the comparative RNA sequencing analyses, and O.G., supervised by M.S., grew fibroblasts, extracted RNA and performed the western blot analyses. N.E., M.T., M.D. and O.E. provided genetic consultation and clinical evaluation for F1. K.M. and K.O. provided genetic consultation and evaluation for F2. S.P., M.H-W., K.O. and K.T.O. performed sequencing and analysis of trio genome and RNA for F2. R.M. and S.A. coordinated the local clinical study for F3-6. M.Z. provided the data for F4, M.Z. and J.G. for F5. D.P. provided genetic consultation and evaluation for F6. P.M., A.G., J.A. and S.L. provided genetic consultation and clinical evaluation for F7-8. C.T.G and T-L.L. performed sequencing and analyses and coordinated the local clinical study for F7-8.

DECLARATION OF INTERESTS

HMS is an employee of Geneyx Genomics. Other authors declare no competing interests.

DATA AVAILABILITY

The ClinVar accession number for the DNA variants data are:

SCV003922044 *WBP4* NM_007187.5 c.499del

SCV003922045 *WBP4* NC_000013.11 g.41074134_41090164del

SCV003922046 *WBP4* NM_007187.5 c.668C>G

SCV003922047 *WBP4* NM_007187.5 c.944del

SCV003922048 *WBP4* NM_007187.5 c.440-1G>A.

All analyzed data consists of patients' personal data and is stored according to regulations of the institutions. Anonymized data is available on request.

REFERENCES

1. Griffin, C., and Saint-Jeannet, J.-P. (2020). Spliceosomopathies: Diseases and mechanisms. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *249*, 1038–1046. 10.1002/dvdy.214.

2. Zhao, C., Bellur, D.L., Lu, S., Zhao, F., Grassi, M.A., Bowne, S.J., Sullivan, L.S., Daiger, S.P., Chen, L.J., Pang, C.P., et al. (2009). Autosomal-dominant retinitis pigmentosa caused by a mutation in SNRNP200, a gene required for unwinding of U4/U6 snRNAs. Am. J. Hum. Genet. *85*, 617–627. 10.1016/j.ajhg.2009.09.020.

3. Knill, C., Henderson, E.J., Johnson, C., Wah, V.Y., Cheng, K., Forster, A.J., and Itasaki, N. (2023). Defects of the spliceosomal gene SNRPB affect osteo- and chondro-differentiation. FEBS J. 10.1111/febs.16934.

4. Bedford, M.T., Reed, R., and Leder, P. (1998). WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. Proc. Natl. Acad. Sci. U. S. A. *95*, 10602–10607. 10.1073/pnas.95.18.10602.

5. Chan, D.C., Bedford, M.T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. EMBO J. *15*, 1045–1054.

6. Huang, X., Beullens, M., Zhang, J., Zhou, Y., Nicolaescu, E., Lesage, B., Hu, Q., Wu, J., Bollen, M., and Shi, Y. (2009). Structure and Function of the Two Tandem WW Domains of the Pre-mRNA Splicing Factor FBP21 (Formin-binding Protein 21). J. Biol. Chem. *284*, 25375–25387. 10.1074/jbc.M109.024828.

7. Klippel, S., Wieczorek, M., Schümann, M., Krause, E., Marg, B., Seidel, T., Meyer, T., Knapp, E.-W., and Freund, C. (2011). Multivalent Binding of Formin-binding Protein 21 (FBP21)-Tandem-WW Domains Fosters Protein Recognition in the Pre-spliceosome. J. Biol. Chem. *286*, 38478–38487. 10.1074/jbc.M111.265710.

8. Henning, L.M., Santos, K.F., Sticht, J., Jehle, S., Lee, C.-T., Wittwer, M., Urlaub, H., Stelzl, U., Wahl, M.C., and Freund, C. (2017). A new role for FBP21 as regulator of Brr2 helicase activity. Nucleic Acids Res. *45*, 7922–7937. 10.1093/nar/gkx535.

9. Bergfort, A., Hilal, T., Kuropka, B., Ilik, İ.A., Weber, G., Aktaş, T., Freund, C., and Wahl, M.C. (2022). The intrinsically disordered TSSC4 protein acts as a helicase inhibitor, placeholder and multi-interaction coordinator during snRNP assembly and recycling. Nucleic Acids Res. *50*, 2938–2958. 10.1093/nar/gkac087.

10. Raghunathan, P.L., and Guthrie, C. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Curr. Biol. CB *8*, 847–855. 10.1016/s0960-9822(07)00345-4.

11. Laggerbauer, B., Achsel, T., and Lührmann, R. (1998). The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices in vitro. Proc. Natl. Acad. Sci. U. S. A. *95*, 4188–4192.

12. Woolard, J., Vousden, W., Moss, S.J., Krishnakumar, A., Gammons, M.V., Nowak, D.G., Dixon, N., Micklefield, J., Spannhoff, A., Bedford, M.T., et al. (2011). Borrelidin modulates the alternative splicing of VEGF in favour of anti-angiogenic isoforms. Chem. Sci. R. Soc. Chem. 2010 *2011*, 273–278. 10.1039/C0SC00297F.

13. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930. 10.1002/humu.22844.

14. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443. 10.1038/s41586-020-2308-7.

15. Lejeune, F. (2022). Nonsense-Mediated mRNA Decay, a Finely Regulated Mechanism. Biomedicines *10*, 141. 10.3390/biomedicines10010141.

16. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. 10.1186/s13059-014-0550-8.

17. Ben-Ari Fuchs, S., Lieder, I., Stelzer, G., Mazor, Y., Buzhor, E., Kaplan, S., Bogoch, Y., Plaschkes, I., Shitrit, A., Rappaport, N., et al. (2016). GeneAnalytics: An Integrative Gene Set Analysis Tool for Next Generation Sequencing, RNAseq and Microarray Data. Omics J. Integr. Biol. *20*, 139–151. 10.1089/omi.2015.0168.

18. Shen, S., Park, J.W., Lu, Z., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U. S. A. *111*, E5593-5601. 10.1073/pnas.1419161111.

19. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., et al. (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585. 10.1038/ng.2653.

20. Magini, P., Smits, D.J., Vandervore, L., Schot, R., Columbaro, M., Kasteleijn, E., van der Ent, M., Palombo, F., Lequin, M.H., Dremmen, M., et al. (2019). Loss of SMPD4 Causes a Developmental Disorder Characterized by Microcephaly and Congenital Arthrogryposis. Am. J. Hum. Genet. *105*, 689–705. 10.1016/j.ajhg.2019.08.006.

21. Grzybowska, E.A. (2012). Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. Biochem. Biophys. Res. Commun. *424*, 1–6. 10.1016/j.bbrc.2012.06.092.

22. Pena, V., Rozov, A., Fabrizio, P., Lührmann, R., and Wahl, M.C. (2008). Structure and function of an RNase H domain at the heart of the spliceosome. EMBO J. *27*, 2929–2940. 10.1038/emboj.2008.209.

23. Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat. Rev. Mol. Cell Biol. *15*, 108–121. 10.1038/nrm3742.

24. Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu. Rev. Biochem. *84*, 291–323. 10.1146/annurev-biochem-060614-034316.

25. Tebbenkamp, A.T.N., Willsey, A.J., State, M.W., and Šestan, N. (2014). The Developmental Transcriptome of the Human Brain: Implications for Neurodevelopmental Disorders. Curr. Opin. Neurol. *27*, 149–156. 10.1097/WCO.0000000000000069.

26. Xu, M., Xie, Y.A., Abouzeid, H., Gordon, C.T., Fiorentino, A., Sun, Z., Lehman, A., Osman, I.S., Dharmat, R., Riveiro-Alvarez, R., et al. (2017). Mutations in the Spliceosome Component CWC27 Cause Retinal Degeneration with or without Additional Developmental Anomalies. Am. J. Hum. Genet. *100*, 592–604. 10.1016/j.ajhg.2017.02.008.

27. Martin, E.M.M.A., Enriquez, A., Sparrow, D.B., Humphreys, D.T., McInerney-Leo, A.M., Leo, P.J., Duncan, E.L., Iyer, K.R., Greasby, J.A., Ip, E., et al. (2020). Heterozygous loss of WBP11 function causes multiple congenital defects in humans and mice. Hum. Mol. Genet. *29*, 3662–3678. 10.1093/hmg/ddaa258.

FIGURES AND TABLES

Table 1. Summary of the demographic, genetic and clinical features of individuals with *WBP4*-related syndrome. All individuals are homozygous for the variant in *WBP4* and the frequency of characteristics is presented as a ratio of 'persons with the feature' to 'persons with data'.

Characteristic	**Individuals with *WBP4*-related syndrome (n=10)**
Age at last investigation, years (n=8)	
Median	3
Range	1.25 to 10
Sex: Female	7/10
Ethnicity	
Egyptian	4/9
Arab Muslim	2/9
Estonian	1/9
Indian	1/9
Moroccan	1/9
WBP4 variant	
c.499del, p.(Thr167Profs*4)	6/10
c.562+1278_*7251del	1/10
c.668C>G, p.(Ser223Ter)	1/10
c.944del, p.(Pro315Glnfs*55)	1/10
c.440-1G>A	1/10
Hypotonia	9/9
Motor delay	8/8
Speech delay	8/8
Dysmorphic facial features	
Abnormality of the outer ear	8/9
Hypertelorism	6/8
Abnormal palate morphology	4/8
Fine hair	3/7
Feeding difficulties / Failure to thrive	7/8
Neuroimaging	
Aplasia/Hypoplasia of the corpus callosum	5/8
Atypical behavior	
Autistic features	5/7
Hyperactivity	5/7
Stereotypy	4/7
Sleep-wake cycle disturbance	4/4
Intellectual disability	5/5
Moderate	1/5
Severe	3/5
Weight and height <-2 SD at last investigation	4/7
Congenital heart defect	4/8
Prenatal issues	4/9
Skeletal abnormalities	3/4
Hearing and/or vision impairments	3/7

Figure 1. Clinical and genetic findings in eight families with *WBP4*-related syndrome. A. Pedigrees of eight families with *WBP4*-related syndrome. Symbols: filled black – affected individuals; diagonal line – deceased; double line – consanguinity; arrows indicate the probands. In each family the likely pathogenic variant in *WBP4* is written beneath the pedigree using NM_007187.5 transcript. Homozygotes for the likely pathogenic variant are noted as -/- and heterozygotes as +/-. F2-II:1 died of aspiration pneumonia at the age of 3 years. F3-II:3 was reported as phenotypically similar to her affected sister, but clinical details were not available. F4-III:1 has Wilson's disease. F5-II:4 died of intestinal obstruction at the age of 55 days; the parents of the proband (I:1 and I:2) are from the same village. F8-II:1 died in infancy due to congenital heart defects. B. Photos of individuals F2-II:1, F4-III:2 and F7-II:1. F2-II:1; profile and frontal photos at the age of 1 day, frontal photo at the age of 5 months and a magnetic resonance image (MRI) at the age of 1.5 years showing agenesis of corpus callosum (arrow). F4-III:2; frontal and profile photos at ~10 years of age and a midsagittal T1-weighted MRI at the age of 1.5 years showing hypoplastic corpus callosum (arrow). F7-II:1; frontal and profile pictures as a newborn and at the age of 4 years and 10 months. C. Schematic diagram of *WBP4* transcript NM_007187.5 and the five LOF variants identified in our study. The gray rectangles represent exons 1-10 starting from left to right. The functional domains (zinc finger, WW1, WW2) of the protein are drawn below as blue and green rectangles. The purple rectangle represents the deletion found in Family 2.

Figure 2. Immunoblotting analysis of WBP4 protein levels in families 1 and 2. A. Western blot was performed to analyze the expression levels of WBP4 and actin proteins in fibroblast cell cultures from control fibroblasts (CCD-1092Sk), Family 1 (F1-II:2 mother, F1-II:1 father, F1-III:1 proband) and Family 2 (F2-I:2 mother, F2-I:1 father, F2-II:1 proband). B. Quantification and relative amounts calculated from the band intensities.

Figure 3. Common abnormal splicing events in F1-III:1 -/- and F2-II:1 -/-. A. RNA was extracted from fibroblasts of F1-II:2 +/-, F1-II:1 +/-, F1-III:1 -/-, F2-I:2 +/-, F2-I:1 +/- and F2-II:1 -/- and subjected to sequencing. Venn diagram representing the overlap in genes with splicing anomalies between the two affected individuals (F1-III:1 and F2-II:1). B. Gene set enrichment analysis for 619 overlapping abnormally spliced genes. Gene set enrichment for human phenotypes, with enrichment represented as the number of genes matching each human phenotype and the enrichment significance (-(Log_{10}) P-value). C-D. Analysis of RNA-seq of probands (F1-III:1 and F2-II:1) relative to available RNA-seq data of age-matched healthy fibroblasts. Summary of significant abnormal splicing events (FDR<0.05, IncLDiff < |0.1|) identified using rMATS. A5SS - alternative 5' splice site; A3SS - alternative 3' splice site, MXE - mutually exclusive exons (C). Gene set enrichment for human phenotypes, with enrichment represented as the number of genes matching each human phenotype and the enrichment significance (-(Log_{10}) P-value) (D).