scientific reports

OPEN

Disease relevance of rare *VPS13B* missense variants for neurodevelopmental Cohen syndrome

Malte Zorn[1,5], Jirko Kühnisch[2,3,5], Sebastian Bachmann[1] & Wenke Seifert[3,4✉]

Autosomal recessive Cohen syndrome is a neurodevelopmental disorder characterized by postnatal microcephaly, intellectual disability, and a typical facial gestalt. Genetic variants in *VPS13B* have been found to cause Cohen syndrome, but have also been linked to autism, retinal disease, primary immunodeficiency, and short stature. While it is well established that loss-of-function mutations of *VPS13B* cause Cohen syndrome, the relevance of missense variants for the pathomechanism remains unexplained. Here, we investigate their pathogenic effect through a systematic re-evaluation of clinical patient information, comprehensive in silico predictions, and in vitro testing of previously published missense variants. In vitro analysis of 10 subcloned *VPS13B* missense variants resulted in full-length proteins after transient overexpression. 6/10 VPS13B missense variants show reduced accumulation at the Golgi complex in the steady state. The overexpression of these 6/10 VPS13B missense variants did not rescue the Golgi fragmentation after the RNAi-mediated depletion of endogenous *VPS13B*. These results thus validate 6/10 missense variants as likely pathogenic according to the classification of the American College of Medical Genetics through the integration of clinical, genetic, in silico, and experimental data. In summary, we state that exact variant classification should be the first step towards elucidating the pathomechanisms of genetically inherited neuronal diseases.

Neurodevelopmental Cohen syndrome is caused by genetic defects in the *VPS13B* (*COH1*) gene[1,2]. Classification and functional validation of genetic disease variants is a critical challenge in clinical diagnostics and the recent explosion of annotated genetic variants strengthen the need to separate disease-related from abundant variants[3,4]. Although a vast number of prediction approaches is established, a large number of disease proteins remains to be functionally characterized[5]. In this regard, characterization of pioneer proteins without clear domain organization and proper functional analysis, such as VPS13B (COH1) is an important research topic. While Cohen syndrome is majorly caused by *VPS13B* loss-of-function mutation, a rare number of missense variants has been published to cause Cohen syndrome. These rare occurrences of disease-related missense variants in VPS13B may provide further insights on protein structure and function and may therefore facilitate the understanding of disease mechanism including genotype/phenotype correlations.

VPS13B is a large and poorly characterized disease-associated protein that is encoded by the *VPS13B* gene[6,7]. *VPS13B* localizes on chromosome 8q22 and spans a genomic region of about 864 kb comprising 62 exons[8]. The VPS13B protein is ~ 450 kDa in size (NP_689777.3, NP_060360.3) and important for Golgi structure maintenance[6]. VPS13B interacts with RAB6, RAB14, STX6, STX13, and PI3P[7,9]. Genetic variants in *VPS13B* have been linked to the neurodevelopmental disorder Cohen syndrome[2,8,10]. The autosomal recessive (AR) Cohen syndrome is characterized by intellectual disability, developmental delay, microcephaly, a characteristic facial appearance, progressive retinopathy, myopia, and/or neutropenia[2,8,11]. Up-to-date over 200 unique and rare variants in *VPS13B* have been reported to cause Cohen syndrome[1,8]. Most of these *VPS13B* variants, including

[1]Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Functional Anatomy, Berlin, Germany. [2]Experimental and Clinical Research Center (ECRC), a cooperation between the Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association and the Charité-Universitätsmedizin Berlin, Berlin, Germany. [3]Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany. [4]Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Cell Biology and Neurobiology, Charitéplatz 1, Berlin 10117, Germany. [5]These authors contributed equally: Malte Zorn and Jirko Kühnisch. ✉email: wenke.seifert@charite.de

nonsense, splice site, frameshift, and large exon-spanning deletions/insertions, can be classified as disease-causing due to their clearly predictable truncation of the encoded protein. Moreover, previous studies identified that premature stop mutations are most likely functional null-alleles due to nonsense-mediated mRNA decay[6,12]. Although the vast majority of *VPS13B* variants are protein length changing, 29 missense variants are reported in association with clinical conditions such as Cohen syndrome, autism, intellectual disability, retinal disease, primary immunodeficiency disease, and/or short stature[1,2,8,13–31]. However, a systematic pathogenicity assessment of these *VPS13B* missense variants has not been performed and their role for Cohen syndrome or other entities is difficult to interpret.

Previous studies showed that AR disorders are habitually caused by pathogenic variants in genes which result in loss-of-function of the encoded protein with none or little residual function. In contrast autosomal dominant (AD) disorders are caused by pathogenic variants which result equally in either loss-of, altered, or gain-of protein function[32]. Moreover, pathogenic variants most likely cluster in AD disorders to functional key regions of the encoded protein, while in AR disorders nearly no clustering of pathogenic loss-of-function variants occurs[32]. This idea can be extended for AR rare pathogenic missense variants. Here, random clusters may support the identification of critical molecular regions[33,34]. For *VPS13B* no clustering of neither loss-of-function nor missense variants could be recognized so far[8]. At the Eighth International Chorea-Acanthocytosis Symposium it was agreed that characterization of rare human missense variants in VPS13 family members, which result in disease but do not cause loss of the specific VPS13 protein, is an important research topic[35].

The mammalian VPS13 protein family consists of four homologues to yeast Vps13p–VPS13A (CHAC), VPS13B (COH1), VPS13C, and VPS13D (Supplementary Fig. S1)[36]. They can be classified as pioneer proteins, since their cell biological as well as molecular function cannot be concluded from sequence homologies. They have in common their giant size (over 300 kDa) and N- and C-terminal VPS13 homology domains[6,36]. To date, genetic alterations in all four *VPS13* genes have been described as causative for various human neurological disease syndromes. Variants in *VPS13A* cause neurodegenerative Chorea-Acanthocytosis[37,38]. Genetic defects in *VPS13B* lead to the neurodevelopmental Cohen syndrome and were also linked to autism, intellectual disability, retinal disease, primary immunodeficiency disease, and short stature[1,2,13–15,19,29,31]. Mutations in *VPS13C* are a rare cause of neurodegenerative early-onset Parkinson's disease[39,40]. And childhood onset of movement disorders is due to variants in *VPS13D*[41,42]. By sequence homologies searches, domain identification databases SMART (http://smart.embl.de/smart/) and pfam (http://pfam.xfam.org/), and literature comparison, we summarize a Chorein/VPS13 region at the very N-terminus (aa 3–102, pfam12624), a second VPS13-N-terminal region (aa 139–280, pfam 16908), a Vps13-adaptor binding (VAB) domain (formerly known as classical DUF1162 or SHR-binding domain, aa 2603–2702), an extended VAB domain (aa 2715–3363, pfam06650)[43], a VPS13-C-terminal region (aa 3543–3709, pfam16909), and an autophagy-related protein C terminal domain (aa 3713–3816, pfam09333) in common with all other VPS13 proteins (Fig. 1A). In contrast to VPS13A, VPS13C, and VPS13D, VPS13B lacks further VPS13-N-terminal mid repeating regions (Supplementary Fig. S2). VPS13B contains in the 6 VAB repeat regions 4 out of 6 highly conserved VPS13B asparagines (manually curated in consistence with[44]). However, we were not able to predict the PH-domain and the APT1 domain at the VPS13B C-terminus using HHpred (https://toolkit.tuebingen.mpg.de/tools/hhpred)[45] or Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/)[46], which has been studied for yeast Vps13p and VPS13A[47,48]. However, it is highly likely that these domains are conserved in all 4 mammalian VPS13 proteins (Supplementary Fig. S2).

All *VPS13*-associated diseases follow an AR inheritance mode, further emphasizing a loss-of-function disease mechanism. Here, we aim to determine the cell biological consequences of the hitherto known *VPS13B* missense variants and will explore their pathogenic potential.

Materials and methods

Study design.
We searched the literature using PubMed and HCMG thoroughly to summarize all published disease-associated *VPS13B* missense variants. Clinical indication and published phenotypical signs were summarized (Supplementary Table T1). We included all mentioned *VPS13B* missense variants regardless of their final interpretation by the respective authors. All *VPS13B* missense variants were subjected to a systematically in silico analysis. Moreover, we searched the gnomAD database for allele frequencies of these missense variants. 10 out of 29 missense variants were cloned as *VPS13B* full length mutants to investigate their impact on Golgi complex function in vitro. Finally, we reclassified their pathogenicity according to the American Collage of Medical Genetics (ACMG) standards[49].

Variant pathogenicity prediction algorithms and ACMG classification.
To assess the pathogenicity of SNPs in silico, a series of established and publicly accessible prediction algorithms were used: Grantham score[50], PROVEAN[51], PolyPhen-2[52], MutationTaster[53], SIFT[54], MutPred[55], and SNPs&GO[56]. Allele frequency was determined using the gnomAD database[57]. Variant classification according to ACMG criteria is a scoring system of precisely defined terms to estimate the impact of a genetic variant on a certain disease or beneficial health state[3,58]. The following information and ACMG terms were implement for classification of genetic *VPS13B* missense variants: (i) variant frequency in control populations (PM2), (ii) computational in silico prediction of the variant impact (PP3), (iii) functional analysis of the variant (PS3), (iv) segregation of the variant within the family (PP1 supporting), and (v) information about the specificity of the phenotype (PP4). The applied terms are summarized for each variant and result in rating a certain variant as pathogenic (P), likely pathogenic (LP), or variant of unknown significance (VUS).

Reagents and antibodies.
All reagents were obtained from Sigma-Aldrich, Roth or Merck unless otherwise stated. The following commercial antibodies were used: Mouse anti-GM130 (BD Transductions Laborato-
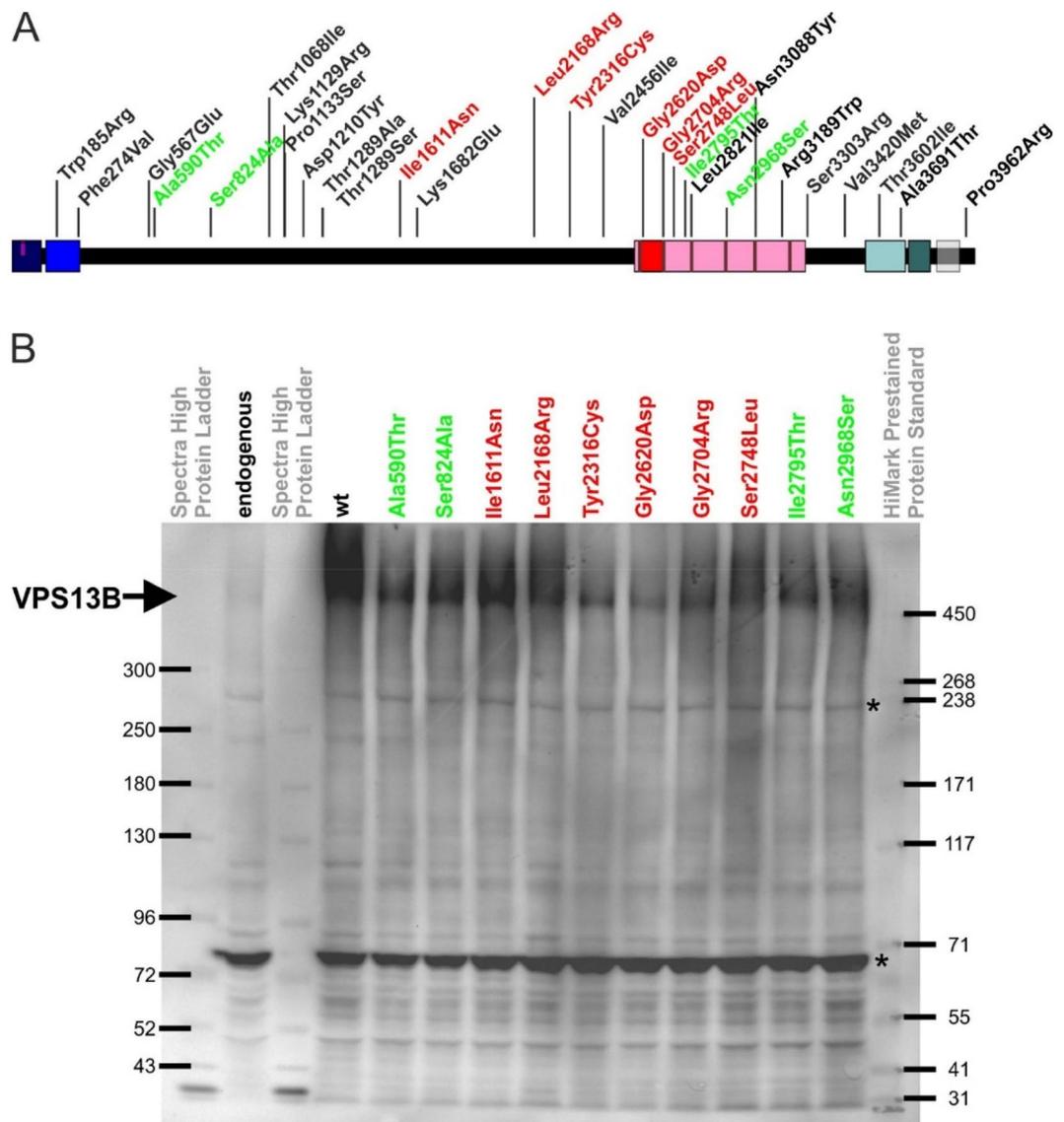
Figure 1. Protein stability of overexpressed wildtype and mutant VPS13B proteins. (**A**) Schematic representation of the VPS13B protein and localization of disease-associated missense variants. Potential functional or homology domains are indicated as follows: Dark blue: a Chorein/VPS13 region at the very N-terminus (aa 3–102, pfam12624), Blue: a second VPS13-N-terminal region (aa 139–280, pfam16908), Red: a Vps13-adaptor binding (VAB) domain (formerly known as classical DUF1162 or SHR-binding domain, aa 2603–2702), Pink: an extended VAB domain (aa 2715–3363, pfam06650)[43], Mint: a VPS13-C-terminal region (aa 3543–3709, pfam16909), and Dark mint: an autophagy-related protein C terminal domain (aa 3713–3816, pfam09333). Color-coded VPS13B missense variants were in vitro analyzed for their Golgi complex association. Missense variants with Golgi association are green while missense variants which fail to associate at the Golgi complex are red. (**B**) HeLa cells were transfected with pcDNA3.1_VPS13B constructs. 24 h post-transfection cells were collected in RIPA lysis buffer, total protein concentration per lane was adjusted and lysates were processed for SDS-PAGE/Western blot analysis by VPS13B. Arrow indicates VPS13B at 450 kDa. *Indicates previously described false positive signals of the VPS13B (442) antibody[6].

ries, 610822, final conc. 2 ng/ml). Rabbit anti-VPS13B (442) was described earlier[6]. Secondary antibodies for Western blotting were anti-rabbit IgG-HRP and anti-mouse IgG-HRP (both DAKO GmbH) and for immuno-fluorescence donkey anti-mouse-Cy3, anti-rabbit-Cy3, and anti-rabbit-Cy2 (all Dianova GmbH). 6-Diamidino-2-phenylindole (DAPI) (Invitrogen) was applied for nuclear DNA-staining.

Cloning of VPS13B mutant expression vectors by site-directed mutagenesis. Genetic variants were assembled from the literature. All missense variants were annotated according to the *VPS13B* (ENST00000357162.6 (*VPS13B*-202) (Ensemble)/NM_152564 (RefSeq)) transcript, including exon 28b in exchange for exon 28 with ubiquitous expression, 14,019 bp transcript length, and encoding 3997 amino acids (aa). For transient expression, full-length human *VPS13B* constructs containing the *VPS13B* missense variants

(pcDNA3.1_VPS13B-mut) were cloned by site-directed mutagenesis PCR[59]. Briefly, full-length wild-type human construct pcDNA3.1_VPS13B-wt[6] was amplified using primer pairs containing the corresponding missense variants in *VPS13B* by Phusion polymerase (NEB Inc.) according to manufacturer's instructions. PCR-products were subsequently treated with endonuclease DpnI (NEB Inc.) to eliminate the methylated pcDNA3.1_VPS13B-wt vectors and further processed for standard DNA ligation reaction. The integrity and correct mutagenesis of all cloned vectors was confirmed by Sanger sequencing.

Cell culture and transfection.   HeLa-cells (ATCC) were cultured in DMEM, 5% FCS, 1% glutamine at 37 °C, 5% CO_2. Transient transfection of expression vectors was performed at 60–75% confluence using jetPEI (Polyplus transfection) according to manufacturer's instructions. Cell harvesting for further analysis was performed 24–48 h after transfection. RNA interference was performed using *VPS13B*-specific siRNA (Ambion); sequences are available on request. The mixture of siRNA, INTERFERin (Polyplus transfection), and OptiMEM (ThermoFisher) was transfected to the cells twice in 12 h intervals according to manufacturer's protocol. The cells were harvested 72 h after the first transfection.

Quantitative PCR.   Isolation of mRNA from HeLa cell-lysates was performed with Isol-RNA Lysis Reagent (5 Prime) according to manufacturer's manual, followed by cDNA synthesis using the M-LV reverse transcriptase (Promega). qPCR analysis was conducted on Fast Real-Time PCR 7500 System (Applied Biosystems) according to the $2^{-\Delta\Delta CT}$ method[60] using the HOT FIREPol EvaGreen HRM (Solis BioDyne) and *VPS13B*-specific primers (sequences available on request). All samples were determined in triplets and normalized against GAPDH expression. Statistical analysis was done by SDS2.1 software (Applied Biosystems).

Immunofluorescence.   For visualization of intracellular protein expression, cells were cultivated on coverslips and if required transfected (see above). Cells on coverslips were fixed using 3% (w/v) paraformaldehyde in PBS at 4 °C and permeabilized in 0,5% (v/v) Triton X-100,1% (w/v) BSA in PBS. Primary antibodies in 1% (w/v) BSA in PBS were incubated for at least 8 h at 4 °C. After rinsing with PBS, secondary antibodies in 1% (w/v) BSA in PBS were incubated for 2 h at 4 °C. Subsequently, coverslips were attached to slides using Fluoromout G (SoutherBiotech). Visualization occurred on the LSM 5 EXCITER confocal microscope and images were analyzed using the ImageJ-software and according macros.

SDS-PAGE/western blotting.   All samples were boiled in $1 \times$ Laemmli buffer and subsequently separated by gel electrophoresis in NuPAGE 3–8% Tris–acetate SDS polyacrylamide gradient gels (Invitrogen) together with Spectra Multicolour High Range Protein Ladder (Thermo Fisher) and HiMark Prestained Protein Standard (life technologies) as range indicators. Separated proteins were transferred to nitrocellulose membranes. After blocking with 5% block milk, membranes were incubated with corresponding primary and HRP-bound secondary antibodies in 5% block milk. Visualization was performed using ECL reaction (Western Blotting Luminol, Santa Cruz Biotechnology) on X-Ray Film (Fuji Medical Super RX).

ImageJ and statistical analysis.   Golgi enrichment was assessed using ImageJ. Two different regions of interest (ROI) were defined for each cell separately: one was drawn around the outline of the cell to detect whole cell total immunofluorescence of VPS13B and the other was drawn around GM130-positive Golgi structure to detect Golgi-associated immunofluorescence of VPS13B. The percentage of Golgi-associated VPS13B fluorescence per total VPS13B cell fluorescence intensity was calculated. The graphical representation as well as the statistical analysis (Dunnett's multiple comparision test) was carried out using GraphPad Prism Software.

Results

To gain further insight into functional VPS13B regions, we summarized literature-published disease-associated VPS13B missense variants, validated their clinical context, predicted the molecular genetic impact, investigated selected variants experimentally, and finally interpreted their disease relevance according to ACMG (Supplementary Fig. S1).

Systematic in silico analyses of published rare disease-associated missense variants in *VPS13B*.   First, we performed a profound literature search on missense variants in *VPS13B* in cases of Cohen syndrome, autism, intellectual disability, retinal disease, primary immunodeficiency disease, and short stature using Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/) and HGMD (http://www.hgmd.cf.ac.uk/ac/index.php). Overall, 29 *VPS13B* disease-associated missense variants were identified[1,2,8,13–23,25–31,61,62] (Table 1). The positions of these missense variants within the VPS13B protein are shown in Fig. 1A. We referenced all *VPS13B* variants to the evolutionary conserved and ubiquitously expressed transcript [NM_152564.4, ENST00000357162] encoding a protein of 3997 aa. However, the annotation to the longest transcript [NM_017890.4, ENST00000358544], which is specifically expressed in the brain and comprises 4022 aa[8], is also given in Supplementary Table T1.

In order to assess the clinical impact of these 29 *VPS13B* missense variants, the genotype and phenotypical characteristics of these cases were summarized (Supplementary Information and Supplementary table T1). The phenotype of each case was critically reviewed by the literature data to assess the probability of a Cohen syndrome diagnosis (Supplementary Information and Supplementary Table T1). The probability of a Cohen syndrome diagnosis, was estimated by using a score counting the main diagnostic criteria: facial dysmorphism, microcephaly, intellectual disability, myopia/retinopathy, neutropenia, and other Cohen syndrome-supportive features (Supplementary Table T1). Although not obligatory, especially neutropenia and/or retinopathy in combination with

Protein change (NP_689777.3)[a]	dbSNP	ClinVar (clinical significance)	Gnomad (minor allele frequency)	Gnomad (homozygous)	Described phenotype	Golgi localization	Golgi rescue	Applicable ACMG pathogenicity criteria	ACMG classification
Wildtype	–	–	No data	No data	–	Yes	Yes		
Trp185Arg	–	No data	No data	No data	Cohen syndrome	n.d	n.d		
Phe274Val	–	No data	0.0004442	0	Autism	n.d	n.d		
Gly567Glu	rs141046414	Uncertain significance	0.0008347	**2**	Undiagnosed genetic condition	n.d	n.d		
Ala590Thr	**rs140601319**	**Uncertain significance**	0.0001592	**1**	**Cohen syndrome**	**Yes**	**Yes**	**None**	**VUS**
Ser824Ala	**rs149866274**	**Uncertain significance**	0.01056	35	**Autism**	**Yes**	**Yes**	**None**	**VUS**
Thr1068Ile	rs61753722	Benign/likely benign	0.004610	5	Epilepsy	n.d	n.d		
Lys1129Arg	rs61759485	Conflicting interpretations	0.000003981	0	Polymorphism	n.d	n.d		
Pro1133Ser	rs781781537	Likely pathogenic	No data	No data	Undiagnosed genetic condition	n.d	n.d		
Asp1210Tyr	–	No data	0.00009212	0	Retinal disease, CS	n.d	n.d		
Thr1289Ala	rs752808333	Uncertain significance	0.0003788	1	Primary immunodeficiency disease	n.d	n.d		
Thr1289Ser	rs145569846	Conflicting interpretations	no data	**No data**	Cohen syndrome	n.d	n.d		
Ile1611Asn	–	**No data**	0.000003540	0	**Cohen syndrome**	**No**	**No**	**PM2, PP3, PS3, PP1sup, PP4**	**LP**
Lys1682Glu	rs965365158	No data	No data	**No data**	Primary immunodeficiency disease	n.d	n.d		
Leu2168Arg	–	**Pathogenic**	0.000007075	**0**	**Cohen syndrome**	**No**	**No**	**PM2, PP3, PS3, PP4**	**LP**
Tyr2316Cys	**rs386834104**	**Likely pathogenic**	0.0001768	0	**Cohen syndrome**	**No**	**No**	**PM2, PP3, PS3, PP4**	**LP**
Val2456Ile	rs201963516	Conflicting interpretations	no data	**No data**	Intellectual disability	n.d	n.d		
Gly2620Asp[b]	–	**Conflicting interpretations**	0.000007975	**0**	**Cohen syndrome**	**No**	**No**	**PM2, PP3, PS3, PP1sup, PP4**	**LP**
Gly2704Arg	**rs767536787**	**No data**	0.000003985	**0**	**Cohen syndrome/Autism**	**No**	**No**	**PM2, PP3, PS3**	**LP**
Ser2748Leu	**rs180177370**	**No data**	No data	**No data**	**Cohen syndrome**	**No**	**No**	**PM2, PP3, PS3, PP1sup, PP4**	**LP**
Ile2795Thr	–	**Pathogenic**	No data	No data	**Cohen syndrome**	**Yes**	**Yes**	**PM2, PP4***	**VUS**
Leu2821Ile		No data	0.003096	**1**	n/p	n.d	n.d		
Asn2968Ser	**rs28940272**	**Conflicting interpretations**	no data	No data	**Cohen syndrome**	**Yes**	**Yes**	**PP3, PP1sup**	**VUS**
Asn3088Tyr	–	No data	0.003778	4	Short stature	n.d	n.d		
Arg3198Trp	rs149842139	Benign/likely benign	no data	No data	Autism	n.d	n.d		
Ser3303Arg	–	No data	0.000202	0	Autism/ Cohen syndrome	n.d	n.d		
Val3445Met	rs191174682	Benign/likely benign	0.000007960	0	Cohen syndrome	n.d	n.d		
Thr3602Ile	–	No data	0.002538	5	Autism/Cohen syndrome	n.d	n.d		
Ala3691Thr	rs142476821	Conflicting interpretations	No data	No data	Autism/ Cohen syndrome	n.d	n.d		
Pro3962Arg	–	Uncertain significance	No data		Autism	n.d	n.d		

Table 1. Summary of results from in vitro characterization of the hitherto cloned VPS13B missense variations and update on ACMG classification. Missense variants in bold were cloned and tested for Golgi localization/rescue and further reclassified according to ACMG guidelines. [a]Conservation of the affected amino acids is given in Supplementary Table T2. [b]Missense variant in combination with consanguineous homozygous loss-of-function variant; *n.d.* not determined.

intellectual disability and microcephaly are clearly pointing toward a Cohen syndrome diagnosis. Accordingly, 19 out of 35 cases demonstrate with a Cohen syndrome consenting phenotype. Several patients from one family were summarized as one case and at least 2 out of 6 criteria (criterium needs to be =/> 50%) have to be extracted from the literature. In 13 out of 35 cases the score for Cohen syndrome was not determinable due to the lack of informative phenotype data. Cases from this group are mostly from different large-scale next generation sequencing (NGS)-screens focusing on one specific clinical sign such as autism, short stature, retinal disease, primary immunodeficiency, and undiagnosed genetic conditions independent of further syndromic features. In three out of 35 cases the Cohen syndrome phenotype score for was < 50%. In this group are: 1 patient with autism, 1 patient with epilepsy and a compound heterozygous so far not confirmed splice site variant c.8016+7G>C, and 1 patient with phenotypically suspected Cohen syndrome but only one heterozygous missense variant. However, two further cases are healthy carriers of one heterozygous missense variant.

All 29 missense variants were evaluated using in silico analyses to predict potential pathogenicity at protein level. These approaches followed three types: (1) sequence and evolutionary conservation-based method, (2) protein sequence and structure-based methods, and (3) supervised machine learning methods. Moreover, the minor allele frequency (MAF) was extracted from GnomAD browser[63]. The accumulated in silico results suggest 12 *VPS13B* missense variants with very low MAF and high disease probability by in silico prediction (Trp185Arg, Asp1210Tyr, Ile1611Asn, Lys1682Glu, Leu2168Arg, Tyr2316Cys, Gly2620Asp, Gly2704Arg, Ser2748Leu, Asn-3088Tyr, Ser3303Arg, and Thr3602Ile) as potentially disease-causing (Table 1, Supplementary Table T2). In addition, 10 *VPS13B* variants were identified with low disease-causing potential by high MAF and negative in silico prediction (Gly567Glu, Ala590Thr, Ser824Ala, Thr1068Ile, Lys1129Arg, Thr1289Ser, Asn2968Ser, Arg3198Trp, Val3420Met, and Ala3691Thr).

Protein levels of full-length VPS13B remained unaffected by missense variants.

The selection on *VPS13B*-associated missense variants for in vitro studies was based on a clear Cohen syndrome/autism association, family-based validation of the genetic finding, and low MAF (Fig. 1A, Table 1). Eight VPS13B missense variants were selected for further in vitro analysis (encoding Ala590Thr, Ile1611Asn, Leu2168Arg, Tyr2316Cys, Gly2620Asp, Ser2748Leu, Ile2795Thr, and Asn2968Ser). These were chosen because prior to the introduction of NGS, larger intragenic deletions and duplications, which are also a common cause of Cohen syndrome, were often overlooked on the second allele[64]. The selection further comprises two missense variants (encoding Ser824Ala and Gly2704Arg), which were identified in a NGS-based screening on autism[31]. To assess the functional impact of these disease-associated missense variants, we subcloned the hitherto selected 10 missense variants into the pcDNA3.1_VPS13B full-length-wild-type (VPS13B-wt) expression vector[6]. To test, whether all VPS13B-mut vectors express normal protein levels, we performed transient overexpression experiments in HeLa cells. The VPS13B-wt as well as all VPS13B-mut proteins are detectable at 450 kDa with similar expression levels suggesting that none of the missense variants cause protein instability (Fig. 1B). None of the mutant VPS13B proteins showed abnormalities with regard to protein degradation or altered general expression.

Golgi fragmentation is rescued by co-overexpression of wildtype-VPS13B.

Immunohistochemical staining and confocal microscopy of transiently transfected HeLa cells revealed that endogenous as well as overexpressed VPS13B-wt co-localized with the Golgi complex marker (GM130) (Fig. 2A,B). This is consistent with previous observations[6]. To exclude possible off-target effects of the hitherto used *VPS13B*-specific siRNA, we established a rescue approach. This includes after *VPS13B* knockdown subsequent transient transfection with a VPS13B-wt expression vector (pcDNA3.1_VPS13B-wt). Effective knockdown of *VPS13B* was confirmed by qPCR (Supplementary Fig. S3) and absence of an appropriate VPS13B protein signal at GM130 positive structures (Fig. 2C). The resulting VPS13B depletion led to severe Golgi fragmentation (Fig. 2C). The Golgi fragmentation and co-localization with GM130 was restored upon by overexpression of VPS13B-wt (Fig. 2D).

Six VPS13B missense variants do not localize to the Golgi complex.

VPS13B is critical for structural integrity of the Golgi complex. VPS13B-deficiency has been shown to fragment Golgi cisternae in fibroblasts derived from Cohen syndrome patients and in *VPS13B*-siRNA knockdown experiments[6]. To assess the subcellular localization of the VPS13B-mut proteins, expression vectors with the corresponding VPS13B missense variant were transfected into Hela cells and the subcellular localization was assessed with immunohistochemical staining, confocal imaging and, fluorescence intensity measurement. Four VPS13B-mut proteins—Ala590Thr, Ser824Ala, Ile2795Thr, and Asn2968Ser—localized to the Golgi complex and showed a similar staining pattern as VPS13B-wt. However, six missense variants—Ile1611Asn, Leu2168Arg, Tyr2316Cys, Gly2620Asp, Gly2704Arg, and Ser2748Leu—did not co-localize with GM130 but demonstrated with a diffuse cytoplasmic staining pattern (Fig. 2E,F, and Supplementary Fig. S4). The significance was calculated using the Dunnett's multiple comparison test GraphPad prism and details of the analysis are provided in the supplement (Supplementary Table T3).

Six VPS13B missense variants do not rescue siRNA-induced Golgi fragmentation.

To further validate the functional overlap of VPS13B-mut compared to VPS13B-wt, we performed a rescue experiment. HeLa cells were treated with a *VPS13B*-specific siRNA inducing fragmentation of the Golgi complex and subsequently transfected with VPS13B-mut vectors. Golgi integrity was assessed with immunohistochemical staining as well as confocal imaging. VPS13B-wt recovered the compact Golgi structure comparable to untreated controls (Fig. 2D). An equivalent effect was observed in four missense variants—Ala590Thr, Ser824Ala, Ile2795Thr, and Asn2968Ser (Fig. 3). In contrast, the other six VPS13B-mut vectors—Ile1611Asn, Leu2168Arg, Tyr2316Cys,
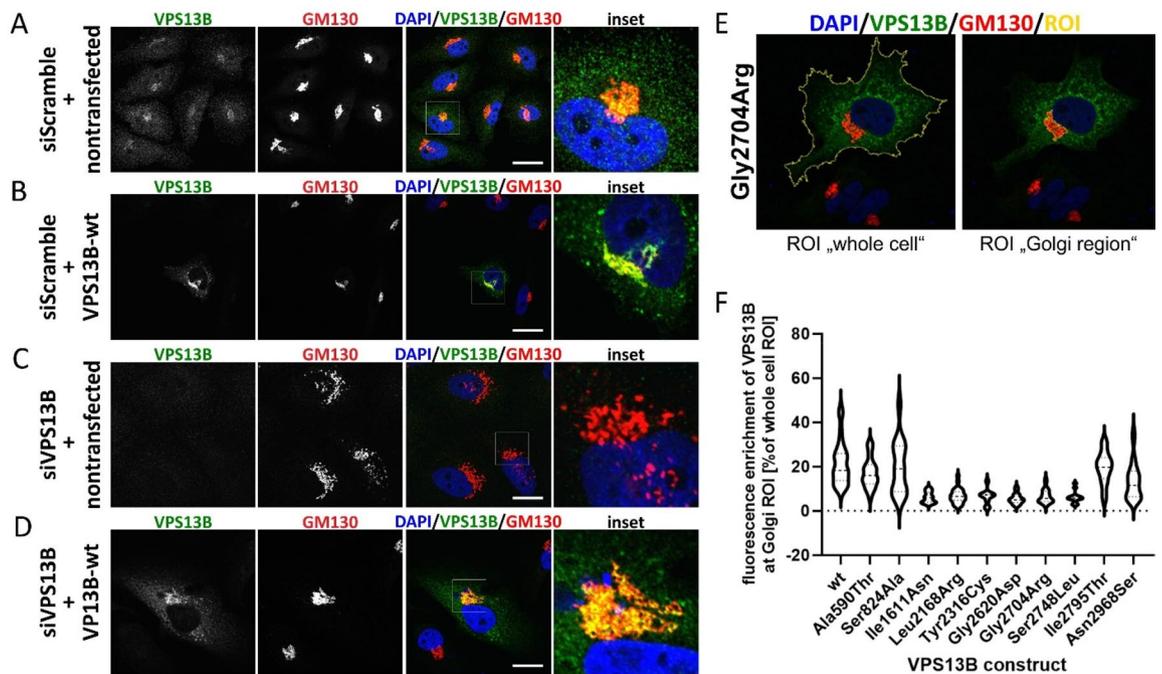
Figure 2. Golgi enrichment is affected by VPS13B missense variants. (**A–D**) Golgi fragmentation is rescued by wild type VPS13B overexpression. HeLa cells were transfected with either control (scrambled siRNA) (**A,B**) or a *VPS13B*-specific RNAi (siRNA #25721, ambion) (**C,D**). 48 h post RNAi-transfection HeLa cells were processed for overexpression and either nontransfected (**A,C**) or transfected with the pcDNA3.1_VPS13B-wt construct (**B,D**). 24 h later, all cells were processed for immunostaining of VPS13B (green) and GM130 (red). Imaging occurred with confocal microscopy. Nuclei were stained using DAPI, scale bars 10 μm. (**E,F**) Analysis of subcellular localization of cloned VPS13B missense variants. ImageJ analysis was used to evaluate Golgi localization of VPS13B missense variants. Two ROI (yellow) were defined for each cell: "whole cell" ROI was drawn at the outline of the cell to detect total immuofluorescence levels; "Golgi region" was defined by the counterstaining of GM130 (red). Fluorescence intensity of VPS13B (green) in each ROI was measured (**E**). Percentage of VPS13B enrichment at the Golgi was finally calculated for each missense variant using GraphPad prism. We repeated each condition at least 3 times and analyzed each time 3 recordings (with at least 1 transfected cell). Please refer to Supplementary table T3 for statistical information (**F**).

Gly2620Asp, Gly2704Arg, and Ser2748Leu—did not restore the compact Golgi morphology and resulted in a diffuse cytoplasmic VPS13B staining.

Six VPS13B missense variants are classified as likely pathogenic according to ACMG guidelines.
For ACMG variant classification we applied input data from multiple levels: MAF from gnomAD population data (PM2), delirious impact of a variant according to computational evidence (PP3), experimental/functional data (PS3), segregation information (PP1), and gene specific phenotype (PP4) (Supplementary Table T2)[3].

Six out of the 10 experimentally tested *VPS13B* missense variants were classified as LP variants (Table 1). The remaining 4 experimentally analyzed variants were classified as VUS. The strongest applied terms for LP classification were very low MAF from gnomAD population data (PM2) and negative functional impact of VPS13B-mut on Golgi organization from in vitro analysis (PS3). It is important to note that 5 *VPS13B* LP variants were detected in patients with a clear diagnosis of Cohen syndrome and in 1 patient with autism. Moreover, the 6 *VPS13B* LP variants were genetically detected in homozygous state in consanguine families or in compound heterozygous state with a *VPS13B* truncating variant.

Discussion
Missense variants are part of the genetic variability within the healthy population, but some usually very rare variants are responsible for genetically inherited diseases. Classifying of rare missense variants with disease states is difficult but provides association with precise genetic defects. Here we assessed the relevance of rare *VPS13B* missense variants for the development of Cohen syndrome. An improved understanding of *VPS13B* missense variants in Cohen syndrome and the affected cellular processes are essential to provide mechanistic explanations and will help to elucidate the disease pathomechanisms.

The members of the VPS13 protein family are considered as pioneering proteins exposing little structural or sequence homologies to other proteins. Recent studies using TAP-tag purified yeast Vps13p and single-particle electron microscopy identified an elongated rod-like structure (~4×20 nm) with a loop (internal diameter ~6 nm) at the one and a hook at the opposite end of yeast Vps13p[70]. Further studies on a N-terminal fragment
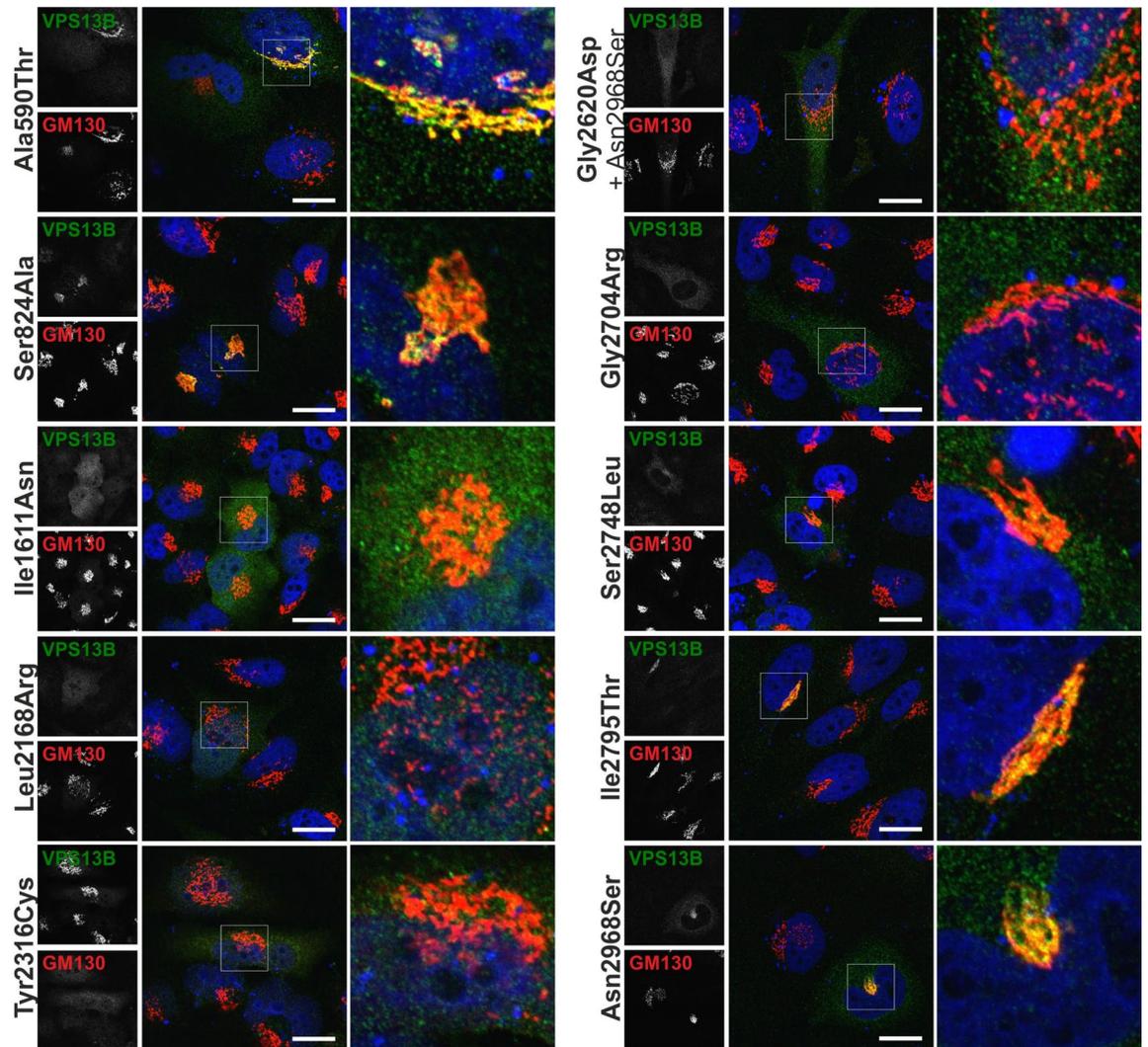
Figure 3. Rescue experiment of Golgi fragmentation by co-overexpression of VPS13B missense variants. HeLa cells were transfected with a *VPS13B*-specific RNAi (siRNA #25721, ambion). 48 h post RNAi-transfection HeLa cells were transfected with the respective pcDNA3.1_VPS13B-mut constructs. 24 h later, cells were processed for immunostaining of VPS13B (green) and GM130 (red). Imaging occurred with confocal microscopy. Nuclei were stained using DAPI, scale bars 10 μm.

from *Cheatomium thermophilum* VPS13 by single particle cryo-EM indicate that the N-terminus is composed of beta-strands forming a long channel-like structure facilitating lipid transport between membranes[65].

Here, six disease-associated VPS13B missense variants were found to disrupt the Golgi complex localization of the VSP13B protein. Moreover, three of these VPS13B missense variants localize in or rather in close proximity with the VAB domain (Gly2620Arg, Gly2704Arg, Ser2748Leu). Although, the function of the VAB domain is not fully understood, recent data suggest that the VAB domain mediates interaction to phosphatidylinositol 3-phosphate (PI3P) in yeast Vps13[66]. Moreover, the VPS13A/CHAC missense variant Ile2771Arg attenuates interaction with PI3P suggesting a role of the VAB domain for direct membrane tethering. Mapping of three VPS13B missense variants to the VAB domain is remarkable as it suggests a similar role of this region for VPS13B. However, limited experimental protein structure information and in silico data prediction hinders assessment of the other three VPS13B missense variants Ile1611Asn, Leu2168Arg, and Tyr2316Cys for overall protein function.

Ongoing utilization of missense variants for functional exploration of VPS13B requires analysis of their disease-causing potential. This analysis involves genetic information, clinical data, in silico predictions, and experimental evidence. Here, we show that functional analysis of disease-associated *VPS13B* missense variants at the Golgi complex critically contributes to variant classification. In this regard, we demonstrate that 6 out of 10 experimentally analyzed *VPS13B* missense variants can be classified as LP according to ACMG. Strongest rating was contributed by the experimental data (PS3) and the low MAF in gnomAD (PM2) stressing the importance of functional read out approaches and population frequency databases for genetic variant interpretation. Together with this, the other four experimentally tested variants are classified as VUS so far. All these variants showed higher MAF values, weak in silico prediction strength, and/or negative experimental outcome suggesting them

as variants without or uncertain functional impact. Further in vitro studies are clearly required to confirm or refute their pathogenicity.

Currently, genetic and phenotype databases foster the implementation of ACMG classified genetic variants. ClinVar is a free accessible archive of clinically relevant variants, which has received an increasing number of submission due to routine genetic testing with NGS[67]. When searching for *VPS13B*, ClinVar shows more than 1000 listed variants. Out of those, 877 *VPS13B* missense variants are listed, most of them as VUS (788), five as LP, and three as P. Eight out of the 29 hitherto summarized missense variants are not listed within ClinVar (Supplementary Table T2). Two out of 29 *VPS13B* missense variants are listed as P (Ile1611Asn, Leu2168Arg) and one variant as LP (Pro1133Ser) (Table 1, https://www.ncbi.nlm.nih.gov/clinvar). Here we confirmed the variants Ile1611Asn and Leu2168Arg as P. Moreover, 4 missense variants (Tyr2316Cys, Gly2620Asp, Gly2704Arg, and Ser2748Leu), which are listed in ClinVar with conflicting interpretations or as VUS, can be classified as LP due to our in vitro study. The provided classification for the *VPS13B* missense variant Pro1133Ser as LP in ClinVar should be recategorized as VUS according to the ACMG guidelines[18]. All *VPS13B* missense variants of this study with a functional impact on Golgi integrity were identified in patients with a definite diagnosis of Cohen syndrome. No genotype–phenotype correlations can currently be performed for Cohen syndrome or *VPS13B*-associated autism. Further research will identify missense variants affecting other VPS13B protein functions and may describe different VPS13B-associated phenotypes in more detail.In conclusion, our study emphasizes that careful and conservative ACMG evaluation as well as consequent database curation is required to avoid mis-, over-, or underinterpretation of genetic *VPS13B* missense variants. This impacts a broad spectrum of diseases and genes in medical genetics[58,68,69]. Consequently, careful investigation and classification of VPS13B missense variants will help to solicit genetic counseling, medical management, and psychosocial guidance of patients as well as their families.

Altogether, our study provides first in vitro evidence that rare *VPS13B* missense variants play a critical role for the pathogenesis of Cohen syndrome. These missense variants compromise, at least partially, the VPS13B protein function. Utilization of further disease-associated missense variants and additional in vitro studies will fundamentally support functional assessment of the VPS13 protein family members and foster classification of VPS13B missense variants, respectively.

References

1. Hennies, H. C. *et al.* Allelic heterogeneity in the COH1 gene explains clinical variability in Cohen syndrome. *Am. J. Hum. Genet.* **75**, 138–145. https://doi.org/10.1086/422219 (2004).
2. Kolehmainen, J. *et al.* Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *Am. J. Hum. Genet.* **72**, 1359–1369. https://doi.org/10.1086/375454 (2003).
3. Richards, 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.* **17**, 405–424. https://doi.org/10.1038/gim.2015.30 (2015).
4. Young, D. L. & Fields, S. The role of functional data in interpreting the effects of genetic variation. *Mol. Biol. Cell* **26**, 3904–3908. https://doi.org/10.1091/mbc.E15-03-0153 (2015).
5. Rodenburg, R. J. The functional genomics laboratory: Functional validation of genetic variants. *J. Inherit. Metab. Dis.* **41**, 297–307. https://doi.org/10.1007/s10545-018-0146-7 (2018).
6. Seifert, W. *et al.* Cohen syndrome-associated protein, COH1, is a novel, giant Golgi matrix protein required for Golgi integrity. *J. Biol. Chem.* **286**, 37665–37675. https://doi.org/10.1074/jbc.M111.267971 (2011).
7. Seifert, W. *et al.* Cohen syndrome-associated protein COH1 physically and functionally interacts with the small GTPase RAB6 at the Golgi complex and directs neurite outgrowth. *J. Biol. Chem.* **290**, 3349–3358. https://doi.org/10.1074/jbc.M114.608174 (2015).
8. Seifert, W. *et al.* Expanded mutational spectrum in Cohen syndrome, tissue expression, and transcript variants of COH1. *Hum. Mutat.* **30**, E404-420. https://doi.org/10.1002/humu.20886 (2009).
9. Koike, S. & Jahn, R. SNAREs define targeting specificity of trafficking vesicles by combinatorial interaction with tethering factors. *Nat. Commun.* **10**, 1608. https://doi.org/10.1038/s41467-019-09617-9 (2019).
10. Momtazmanesh, S. *et al.* A novel VPS13B mutation in Cohen syndrome: A case report and review of literature. *BMC Med. Genet.* **21**, 140. https://doi.org/10.1186/s12881-020-01075-1 (2020).
11. Wang, H., Falk, M. J., Wensel, C., Traboulsi, E. I. *Cohen Syndrome.* (eds Adam, M. P., *et al.*) (GeneReviews((R)), 1993).
12. Boschann, F. *et al.* An intronic splice site alteration in combination with a large deletion affecting VPS13B (COH1) causes Cohen syndrome. *Eur. J. Med. Genet.* **63**, 103973. https://doi.org/10.1016/j.ejmg.2020.103973 (2020).
13. Carss, K. J. *et al.* Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. *Am. J. Hum. Genet.* **100**, 75–90. https://doi.org/10.1016/j.ajhg.2016.12.003 (2017).
14. Chi, Z. H. *et al.* Targeted high-throughput sequencing technique for the molecular diagnosis of primary immunodeficiency disorders. *Medicine (Baltimore)* **97**, e12695. https://doi.org/10.1097/MD.0000000000012695 (2018).
15. de Ligt, J. *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* **367**, 1921–1929. https://doi.org/10.1056/NEJMoa1206524 (2012).
16. El Chehadeh, S. *et al.* Search for the best indicators for the presence of a VPS13B gene mutation and confirmation of diagnostic criteria in a series of 34 patients genotyped for suspected Cohen syndrome. *J. Med. Genet.* **47**, 549–553. https://doi.org/10.1136/jmg.2009.075028 (2010).
17. Falk, M. J. *et al.* Cohen syndrome in the Ohio Amish. *Am. J. Med. Genet. A* **128A**, 23–28. https://doi.org/10.1002/ajmg.a.30033 (2004).
18. Farwell, K. D. *et al.* Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: Results from 500 unselected families with undiagnosed genetic conditions. *Genet. Med.* **17**, 578–586. https://doi.org/10.1038/gim.2014.154 (2015).
19. Huang, Z. *et al.* Genetic evaluation of 114 Chinese short stature children in the next generation era: A single center study. *Cell Physiol. Biochem.* **49**, 295–305. https://doi.org/10.1159/000492879 (2018).

20. Ionita-Laza, I., Capanu, M., De Rubeis, S., McCallum, K. & Buxbaum, J. D. Identification of rare causal variants in sequence-based studies: Methods and applications to VPS13B, a gene involved in Cohen syndrome and autism. *PLoS Genet.* **10**, e1004729. https://doi.org/10.1371/journal.pgen.1004729 (2014).

21. Katzaki, E. *et al.* Clinical and molecular characterization of Italian patients affected by Cohen syndrome. *J. Hum. Genet.* **52**, 1011–1017. https://doi.org/10.1007/s10038-007-0208-4 (2007).

22. Kolehmainen, J. *et al.* Delineation of Cohen syndrome following a large-scale genotype–phenotype screen. *Am. J. Hum. Genet.* **75**, 122–127. https://doi.org/10.1086/422197 (2004).

23. Koshimizu, E. *et al.* Performance comparison of bench-top next generation sequencers using microdroplet PCR-based enrichment for targeted sequencing in patients with autism spectrum disorder. *PLoS One* **8**, e74167. https://doi.org/10.1371/journal.pone.0074167 (2013).

24. Lee, Y. K. *et al.* Cohen syndrome patient iPSC-derived neurospheres and forebrain-like glutamatergic neurons reveal reduced proliferation of neural progenitor cells and altered expression of synapse genes. *J. Clin. Med.* https://doi.org/10.3390/jcm9061886 (2020).

25. Mochida, G. H. *et al.* Broader geographical spectrum of Cohen syndrome due to COH1 mutations. *J. Med. Genet.* **41**, e87. https://doi.org/10.1136/jmg.2003.014779 (2004).

26. Retterer, K. *et al.* Clinical application of whole-exome sequencing across clinical indications. *Genet. Med.* **18**, 696–704. https://doi.org/10.1038/gim.2015.148 (2016).

27. Rivera-Brugues, N. *et al.* Cohen syndrome diagnosis using whole genome arrays. *J. Med. Genet.* **48**, 136–140. https://doi.org/10.1136/jmg.2010.082206 (2011).

28. Seifert, W. *et al.* Mutational spectrum of COH1 and clinical heterogeneity in Cohen syndrome. *J. Med. Genet.* **43**, e22. https://doi.org/10.1136/jmg.2005.039867 (2006).

29. Stone, E. M. *et al.* Clinically focused molecular investigation of 1000 consecutive families with inherited retinal disease. *Ophthalmology* **124**, 1314–1331. https://doi.org/10.1016/j.ophtha.2017.04.008 (2017).

30. Yang, C. *et al.* Gene analysis: A rare gene disease of intellectual deficiency-Cohen syndrome. *Int. J. Dev. Neurosci.* **68**, 83–88. https://doi.org/10.1016/j.ijdevneu.2018.05.004 (2018).

31. Yu, T. W. *et al.* Using whole-exome sequencing to identify inherited causes of autism. *Neuron* **77**, 259–273. https://doi.org/10.1016/j.neuron.2012.11.002 (2013).

32. Turner, T. N. *et al.* Proteins linked to autosomal dominant and autosomal recessive disorders harbor characteristic rare missense mutation distribution patterns. *Hum. Mol. Genet.* **24**, 5995–6002. https://doi.org/10.1093/hmg/ddv309 (2015).

33. Schindelhauer, D. *et al.* Wiskott–Aldrich syndrome: No strict genotype-phenotype correlations but clustering of missense mutations in the amino-terminal part of the WASP gene product. *Hum. Genet.* **98**, 68–76. https://doi.org/10.1007/s004390050162 (1996).

34. Weile, J. *et al.* Ultrasonography in trauma: A nation-wide cross-sectional investigation. *Crit. Ultrasound J.* **9**, 16. https://doi.org/10.1186/s13089-017-0071-2 (2017).

35. Pappas, S. S. *et al.* Eighth international chorea-acanthocytosis symposium: Summary of workshop discussion and action points. *Tremor Other Hyperkinet Mov (N Y)* **7**, 428. https://doi.org/10.7916/D8XD127W (2017).

36. Velayos-Baeza, A., Vettori, A., Copley, R. R., Dobson-Stone, C. & Monaco, A. P. Analysis of the human VPS13 gene family. *Genomics* **84**, 536–549. https://doi.org/10.1016/j.ygeno.2004.04.012 (2004).

37. Rampoldi, L. *et al.* A conserved sorting-associated protein is mutant in chorea-acanthocytosis. *Nat. Genet.* **28**, 119–120 (2001).

38. Ueno, S. *et al.* The gene encoding a newly discovered protein, chorein, is mutated in chorea-acanthocytosis. *Nat. Genet.* **28**, 121–122 (2001).

39. Schormair, B. *et al.* Diagnostic exome sequencing in early-onset Parkinson's disease confirms VPS13C as a rare cause of autosomal-recessive Parkinson's disease. *Clin. Genet.* **93**, 603–612. https://doi.org/10.1111/cge.13124 (2018).

40. Zou, M. *et al.* Association analyses of variants of SIPA1L2, MIR4697, GCH1, VPS13C, and DDRGK1 with Parkinson's disease in East Asians. *Neurobiol. Aging* https://doi.org/10.1016/j.neurobiolaging.2018.03.005 (2018).

41. Gauthier, J. *et al.* Recessive mutations in VPS13D cause childhood onset movement disorders. *Ann. Neurol.* https://doi.org/10.1002/ana.25204 (2018).

42. Seong, E. *et al.* Mutations in VPS13D lead to a new recessive ataxia with spasticity and mitochondrial defects. *Ann. Neurol.* https://doi.org/10.1002/ana.25220 (2018).

43. Bean, B. D. M. *et al.* Competitive organelle-specific adaptors recruit Vps13 to membrane contact sites. *J. Cell Biol.* **217**, 3593–3607. https://doi.org/10.1083/jcb.201804111 (2018).

44. Dziurdzik, S. K., Bean, B. D. M., Davey, M. & Conibear, E. A VPS13D spastic ataxia mutation disrupts the conserved adaptor-binding site in yeast Vps13. *Hum. Mol. Genet.* **29**, 635–648. https://doi.org/10.1093/hmg/ddz318 (2020).

45. Zimmermann, L. *et al.* A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007 (2018).

46. Wass, M. N., Kelley, L. A. & Sternberg, M. J. 3DLigandSite: Predicting ligand-binding sites using similar structures. *Nucleic Acids Res.* **38**, W469-473. https://doi.org/10.1093/nar/gkq406 (2010).

47. Fidler, D. R. *et al.* Using HHsearch to tackle proteins of unknown function: A pilot study with PH domains. *Traffic* **17**, 1214–1226. https://doi.org/10.1111/tra.12432 (2016).

48. Kolakowski, D., Kaminska, J. & Zoladek, T. The binding of the APT1 domains to phosphoinositides is regulated by metal ions in vitro. *Biochim. Biophys. Acta Biomembr.* **1862**, 183349. https://doi.org/10.1016/j.bbamem.2020.183349 (2020).

49. Deignan, J. L. *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.* **21**, 1267–1270. https://doi.org/10.1038/s41436-019-0478-1 (2019).

50. Grantham, R. Amino acid difference formula to help explain protein evolution. *Science* **185**, 862–864 (1974).

51. Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**, 2745–2747. https://doi.org/10.1093/bioinformatics/btv195 (2015).

52. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249. https://doi.org/10.1038/nmeth0410-248 (2010).

53. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: Mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362. https://doi.org/10.1038/nmeth.2890 (2014).

54. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081. https://doi.org/10.1038/nprot.2009.86 (2009).

55. Li, B. *et al.* Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* **25**, 2744–2750. https://doi.org/10.1093/bioinformatics/btp528 (2009).

56. Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L. & Casadio, R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum. Mutat.* **30**, 1237–1244. https://doi.org/10.1002/humu.21047 (2009).

57. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291. https://doi.org/10.1038/nature19057 (2016).

58. Amendola, L. M. *et al.* Performance of ACMG-AMP variant-interpretation guidelines among nine laboratories in the clinical sequencing exploratory research consortium. *Am. J. Hum. Genet.* **98**, 1067–1076. https://doi.org/10.1016/j.ajhg.2016.03.024 (2016).

59. Reikofski, J. & Tao, B. Y. Polymerase chain reaction (PCR) techniques for site-directed mutagenesis. *Biotechnol. Adv.* **10**, 535–547 (1992).

60. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).
61. Dopazo, J. *et al.* 267 Spanish exomes reveal population-specific differences in disease-related genetic variation. *Mol. Biol. Evol.* **33**, 1205–1218. https://doi.org/10.1093/molbev/msw005 (2016).
62. Li, A., Gandhi, A., Wang, H. & Traboulsi, E. I. Bilateral angle closure glaucoma in a 28-year-old Cohen syndrome patient. *Ophthalmic Genet.* **39**, 657–658. https://doi.org/10.1080/13816810.2018.1495746 (2018).
63. Karczewski, K. J. *et al.* Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv* https://doi.org/10.1101/531210 (2019).
64. Parri, V. *et al.* High frequency of COH1 intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome. *Eur. J. Hum. Genet.* **18**, 1133–1140. https://doi.org/10.1038/ejhg.2010.59 (2010).
65. Li, P., Lees, J. A., Lusk, C. P. & Reinisch, K. M. Cryo-EM reconstruction of a VPS13 fragment reveals a long groove to channel lipids between membranes. *J. Cell Biol.* https://doi.org/10.1083/jcb.202001161 (2020).
66. Rzepnikowska, W. *et al.* Amino acid substitution equivalent to human chorea-acanthocytosis I2771R in yeast Vps13 protein affects its binding to phosphatidylinositol 3-phosphate. *Hum. Mol. Genet.* **26**, 1497–1510. https://doi.org/10.1093/hmg/ddx054 (2017).
67. Landrum, M. J. & Kattman, B. L. ClinVar at five years: Delivering on the promise. *Hum. Mutat.* **39**, 1623–1630. https://doi.org/10.1002/humu.23641 (2018).
68. Giudicessi, J. R., Roden, D. M., Wilde, A. A. M. & Ackerman, M. J. Classification and reporting of potentially proarrhythmic common genetic variation in long QT syndrome genetic testing. *Circulation* **137**, 619–630. https://doi.org/10.1161/CIRCULATIONAHA.117.030142 (2018).
69. Oldak, M., Lechowicz, U., Pollak, A., Ozieblo, D. & Skarzynski, H. Overinterpretation of high throughput sequencing data in medical genetics: First evidence against TMPRSS3/GJB2 digenic inheritance of hearing loss. *J. Transl. Med.* **17**, 269. https://doi.org/10.1186/s12967-019-2018-9 (2019).
70. De, M. *et al.* (2017) The Vps13p-Cdc31p complex is directly required for TGN late endosome transport and TGN homotypic fusion. *J. Cell.Biol.* **216**, 425–439. https://doi.org/10.1083/jcb.201606078

Acknowledgements

Author contributions

W.S. designed the study; W.S. and J.K. were major contributors in manuscript writing, M.Z. performed literature search; W.S., M.Z. performed cloning and cell biological analyses; S.B. provided expertise and resources. All authors read and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-13717-w.

Correspondence and requests for materials should be addressed to W.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.