REPORT

Mutations in the BAF-Complex Subunit *DPF2* Are Associated with Coffin-Siris Syndrome

Georgia Vasileiou,[1] Silvia Vergarajauregui,[2] Sabine Endele,[1] Bernt Popp,[1] Christian Büttner,[1] Arif B. Ekici,[1] Marion Gerard,[3] Nuria C. Bramswig,[4] Beate Albrecht,[4] Jill Clayton-Smith,[5] Jenny Morton,[6] Susan Tomkins,[7] Karen Low,[7] Astrid Weber,[8] Maren Wenzel,[9] Janine Altmüller,[10] Yun Li,[11] Bernd Wollnik,[11] George Hoganson,[12] Maria-Renée Plona,[12] Megan T. Cho,[13] Deciphering Developmental Disorders Study,[14] Christian T. Thiel,[1] Hermann-Josef Lüdecke,[4,15] Tim M. Strom,[16] Eduardo Calpena,[17] Andrew O.M. Wilkie,[17] Dagmar Wieczorek,[4,15] Felix B. Engel,[2] and André Reis[1,*]

Variants affecting the function of different subunits of the BAF chromatin-remodelling complex lead to various neurodevelopmental syndromes, including Coffin-Siris syndrome. Furthermore, variants in proteins containing PHD fingers, motifs recognizing specific histone tail modifications, have been associated with several neurological and developmental-delay disorders. Here, we report eight heterozygous *de novo* variants (one frameshift, two splice site, and five missense) in the gene encoding the BAF complex subunit double plant homeodomain finger 2 (DPF2). Affected individuals share common clinical features described in individuals with Coffin-Siris syndrome, including coarse facial features, global developmental delay, intellectual disability, speech impairment, and hypoplasia of fingernails and toenails. All variants occur within the highly conserved PHD1 and PHD2 motifs. Moreover, missense variants are situated close to zinc binding sites and are predicted to disrupt these sites. Pull-down assays of recombinant proteins and histone peptides revealed that a subset of the identified missense variants abolish or impaire DPF2 binding to unmodified and modified H3 histone tails. These results suggest an impairment of PHD finger structural integrity and cohesion and most likely an aberrant recognition of histone modifications. Furthermore, the overexpression of these variants in HEK293 and COS7 cell lines was associated with the formation of nuclear aggregates and the recruitment of both wild-type DPF2 and BRG1 to these aggregates. Expression analysis of truncating variants found in the affected individuals indicated that the aberrant transcripts escape nonsense-mediated decay. Altogether, we provide compelling evidence that *de novo* variants in *DPF2* cause Coffin-Siris syndrome and propose a dominant-negative mechanism of pathogenicity.

Coffin-Siris syndrome (CSS [MIM: 135900]) is a neurodevelopmental disorder characterized by mild to severe intellectual disability, speech impairment, growth deficiency, feeding difficulties, coarse facial characteristics, sparse hair, hypoplastic or absent finger- and/or toenails, and brain anomalies, the most prominent of which is hypoplasia or agenesis of the corpus callosum.[1,2] The rather broad and highly variable clinical spectrum of CSS individuals often confounds the clinical diagnosis.[3] In recent years, germline *de novo* mutations in several subunits of the human BRG1-associated factor (BAF) chromatin-remodelling complex (also known as the SWI/SNF-A complex) have been associated with CSS. Heterozygous loss-of-function variants (frameshifts, nonsense variants, and microdeletions) have been identified in *ARID1A* (MIM: 603024) and *ARID1B* (MIM: 614556), whereas variants with dominant-negative or gain-of-function effects (missense and in-frame deletions) have been detected in *SMARCA4* (*BRG1* [MIM: 603254]), *SMARCB1* (*SNF5* [MIM: 601607]), and *SMARCE1* (MIM: 603111).[1,4–11] Furthermore, two groups have reported haploinsufficiency of *SOX11* (MIM: 600898), a downstream transcriptional factor of the BAF complex, in individuals with mild to severe CSS.[10,12] Finally, loss-of-function variants in *ARID2* (MIM: 609539) have been recently associated with a CSS-like phenotype.[13]

Here, we report *de novo* variants in *DPF2* (also known as *REQ*, *UBID4*, or *BAF45d* [MIM: 601671]), encoding a subunit of the BAF chromatin-remodelling complex, which has not previously been associated with neurodevelopmental syndromes. In total, we identified eight unrelated individuals (four females and four males) with features of

[1]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany; [2]Experimental Renal and Cardiovascular Research, Institute of Pathology, Department of Nephropathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91054 Erlangen, Germany; [3]Génétique Clinique, Centre Hospitalier Universitaire de Caen, Caen 14000, France; [4]Institut für Humangenetik, Universitätsklinikum Essen, Universität Duisburg-Essen, 45122 Essen, Germany; [5]Manchester Centre for Genomic Medicine, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester M13 9WL, UK; [6]West Midlands Regional Clinical Genetics Service and Birmingham Health Partners, Birmingham Women's Hospital NHS Foundation Trust, Birmingham B15 2TG, UK; [7]Clinical Genetics Service, University Hospitals of Bristol NHS Foundation Trust, Bristol BS2 8HW, UK; [8]Merseyside and Cheshire Clinical Genetics Service, Liverpool Women's NHS Foundation Hospital Trust, Liverpool L8 7SS, UK; [9]Genetikum Neu-Ulm, 89231 Neu-Ulm, Germany; [10]Cologne Center for Genomics, University of Cologne, 50931 Cologne, Germany; [11]Institute of Human Genetics, University Medical Center Göttingen, 37073 Göttingen, Germany; [12]Pediatric Genetics, University of Illinois Hospital, Chicago, IL 60612, USA; [13]GeneDx, Gaithersburg, MD 20877, USA; [14]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; [15]Institut für Humangenetik, Universitätsklinikum Düsseldorf, Heinrich-Heine-Universität, 40225 Düsseldorf, Germany; [16]Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg 85764, Germany; [17]Clinical Genetics Group, MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK
*Correspondence: andre.reis@uk-erlangen.de
https://doi.org/10.1016/j.ajhg.2018.01.014.

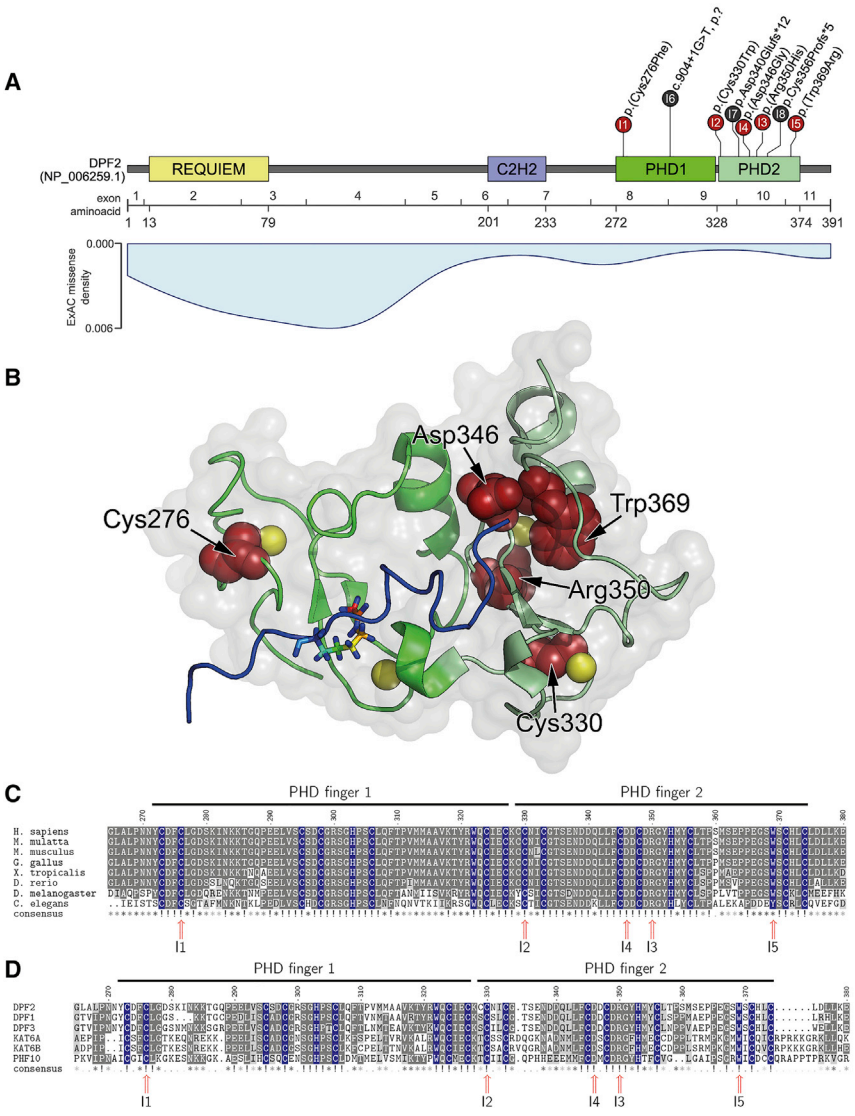

Figure 1. Clustering of *DPF2* Variants in PHD Fingers

(A) Schematic representation of DPF2, its domains (based on GenBank: NP_006259.1), the encoding exons (numbering based on GenBank: NM_006268.4), and the localization of *DPF2* variants. Missense variants are presented in red, and truncating variants are in black. Note that the premature termination codons of the two truncating variants p.Asp340Glufs*12 and p.Cys356Profs*5 reside within 50 nucleotides upstream of the most 3′ exon-exon junction. The numbers in the circles indicate the affected individuals. For individual 6, c.904+1G>T is described at the genomic level because no RT-PCR could be performed. In light blue, a density blot of all missense variants reported in ExAC Browser version 0.3.1 shows markedly low frequency of variants in the PHD zinc fingers.

(B) The crystal structure of the DPF2 double PHD finger bound to a histone peptide containing acetylation at lysine 14 (H3K14ac) (PDB: 5B79[18] and 2KWJ[19]) shows the clustering of the herein described missense variants in the tandem PHD finger domain (PHD finger 1 is colored in bright green, and PHD finger 2 is colored in pale green). The histone H3 backbone in the DPF2 binding pocket is shown in blue with an acetylated lysine residue at position 14 in stick representation. Zinc ions are represented as yellow spheres. The five affected amino acid residues are colored in red. The electrostatic surface is represented in gray with 80% opacity. Cys276 and Asp346 reside at the protein surface, whereas Cys330, Arg350, and Trp369 are buried in the PHD2 domain.[20]

(C) Multiple-sequence alignment of DPF2 orthologs at the *de novo* missense variant positions in the tandem PHD fingers shows high evolutionary sequence conservation. Residues from the conserved C4HC3 signature are marked in blue (see also Figure S5A and Data S1 sheet "DPF2_orthologs"). Numbers I1–I5 indicate the individual with the respective variant. Positions with *de novo* missense variants are indicated with a red arrow. Gray shading represents conservation.

(D) Amino acid sequence alignment of DPF2 and putative human paralog proteins with similar tandem PHD fingers shows conservation of the C4HC3 signature (see also Figure S5B and Data S1 sheet "PHD_finger_proteins_(PHF)"). Protein sequences were obtained from NCBI and ClustalW,[21] and the msa package[22] within R was used for alignment. References for the orthologs are as follows: *H. sapiens* (DPF2), GenBank: NP_006259.1; *M. mulatta* (LOC721967), GenBank: XP_002808108.1; *M. musculus* (Dpf2), GenBank: NP_035392.1; *G. gallus* (DPF2), GenBank: NP_989662.1; *D. rerio* (dpf2), GenBank: NP_001007153.1; and *X. tropicalis* (dpf2), GenBank: NP_001184101.1. References for the paralogs are as follows: DPF1, GenBank: NP_001128627.1; DPF2, GenBank: NP_006259.1; DPF3, GenBank: NP_001267471.1; KAT6A, GenBank: NP_006757.2; KAT6B, GenBank: NP_036462.2; and PHF10, GenBank: NP_060758.2.

CSS. Written informed consent was obtained from all participants or their legal guardians, and the study was approved by the ethical review board of the University Erlangen-Nürnberg and the respective institutions, as well as the East of England-Cambrige South committee of the National Research Ethics Service for the UK Deciphering Developmental Disorders (DDD) Study. The study received UK research ethics committee (REC) approval (10/H0305/83 granted by the Cambridge South REC and GEN/284/12 granted by the Republic of Ireland REC).

DPF2 is located in chromosomal region 11q13.1 and shows a ubiquitous expression pattern. DPF2, together with DPF1 and DPF3, belongs to the d4 protein family and functions as a non-catalytic subunit of the BAF chromatin-remodelling complex.[14–17] The protein contains three domains (Figure 1A): (1) an N-terminal requiem domain, which has been shown to interact with catalytic (SMARCA4 and SMARCA2), core (SMARCB1 and SMARCC1), and other (SMARCD1) subunits of the BAF complex; (2) a single Krüppel-type zinc finger domain (C2H2) specific to DNA binding; and (3) a C-terminal tandem plant homeodomain (PHD) finger (PHD1 and PHD2). PHD zinc fingers mediate protein interactions, bind on nucleosomes, and recognize post-translational histone

modifications.[16,18,23,24] They are highly conserved structures containing a zinc binding motif with a Cys4-His-Cys3 architecture (C4HC3) that anchors two Zn^{2+} ions.[25]

DPF2 was first identified as an early-apoptosis gene in mouse myeloid cells.[26] Recently, its inhibitory role in myeloid differentiation via interaction with RUNX1, possibly mediated by the BAF complex, was reported.[23] Furthermore, DPF2 modulates the association between the BRM-type BAF complex and the protein dimer RelB/p52 to activate the noncanonical NF-kB pathway supporting oncogenesis and promotes the ubiquitination and degradation of OCT4, a crucial stem cell pluripotency factor.[16,27]

The index subject (individual 1), a 10-year-old boy from France, was clinically suspected to have CSS on the basis of his coarse facial features, global developmental delay, mild intellectual disability, speech delay (first words at 48 months and whole sentences at 84 months), stereotypic behavior, feeding problems for 1–2 years starting at the age of 6 months, and muscular hypotonia. The boy was able to walk after the age of 17 months. Hypoplasia of the fourth and fifth toenails and brachydactyly of the fifth fingers were also observed (Figure 2, Table 1, and Table S1). Chromosomal microarray analysis did not reveal any pathogenic copy-number variants (CNVs), and Sanger sequencing of six genes (*SMARCA4*, *SMARCB1*, *SMARCE1*, *ARID1B*, *ARID1A*, and *SMARCA2* [MIM: 600014]) associated with CSS and Nicolaides-Baraitser syndrome (MIM: 601358) showed no variant explaining the observed phenotype (Supplemental Note and Table S2). Subsequent trio exome sequencing on an Illumina HiSeq 2500 system after enrichment with SureSelect Target Enrichment V5 technology (Agilent Technologies) revealed the previously unreported heterozygous *de novo* missense variant c.827G>T (p.Cys276Phe) in exon 8 of *DPF2*. The nucleotide change affects a highly conserved amino acid in the PHD1 finger (Figure 1). The variant has not been described in the Genome Aggregation Database (gnomAD) and is computationally predicted to be deleterious (Table S3).

Additional evidence for the correlation between the identified *DPF2* variant and the clinical phenotype was provided by a large exome-wide trio study on 4,293 children with severe, undiagnosed developmental disorders (DDD Study).[28] From the *de novo* variants reported, we extracted two additional heterozygous *DPF2* missense variants in female individuals originating from England: c.990C>G (p.Cys330Trp) in exon 9 (individual 2 [DDD4K.02108]) and c.1049G>A (p.Arg350His) in exon 10 (individual 3 [DDD4K.03804]). Further interrogation of the DDD resource at a later stage for subjects with craniosynostosis (Datafreeze 3: 7,833 trios and 1,792 singletons) revealed *DPF2* missense mutations in two further individuals, tested by trio exome sequencing: c.1037A>G (p.Asp346Gly) in exon 10 (individual 4 [DDDP126054]) and c.1105T>C (p.Trp369Arg) in exon 11 (individual 5 [DDDP100312]). All four variants are located in the PHD2 finger and affect highly conserved amino acids (Figure 1). They are likewise not present in gnomAD, and the computer-based prediction programmes classified them as deleterious (Table S3).

On the basis of further collaborations and by employing the web-based matching platform GeneMatcher,[29] we identified another three individuals with *de novo DPF2* likely gene-disrupting (LGD) variants, all located in one of the PHD domains. These individuals underwent trio exome sequencing at their respective institutions.[30–32] Premature termination codons occurring at least 50 bp upstream of the last exon-exon junction have been described to escape nonsense-mediated decay (NMD).[33,34] Two of the variants fell under this category. mRNA sequencing for c.1099+1G>A (individual 7), which affects the consensus GT splice-donor site in intron 10, showed a heterozygous aberrant transcript with skipping of exon 10 (r.1018_1099del). This resulted in a frameshift and a premature stop codon after 12 amino acids (p.Asp340Glufs*12) (Figure 1A and Figures S1A und S1B). RNA expression analysis also confirmed the 8 bp, exon 10 heterozygous deletion c.1066_1073del (individual 8), predicted to cause a frameshift with a premature termination codon after five residues (p.Cys356Profs*5), virtually excluding NMD (Figure 1A and Figures S1C and S1D). The third possible LGD variant was the *de novo* splice-donor mutation c.904+1G>T in individual 6 (Figure 1A). An RNA sample from this individual was not available. This variant is located >50 bp upstream of the last exon-exon junction. However, it is predicted to disrupt the affected canonical splice-donor site of intron 8 (Table S3) and most likely lead to the skipping of exon 8 (r.776_904del). The predicted skipping is likely to result in a stable, in-frame insertion or deletion (p.Ser259_Gly302delinsTrp) affecting several highly conserved amino acids of PHD1. Apart from the *DPF2* variants, the broad genetic evaluation revealed additional rare single-nucleotide variants or CNVs in all individuals except individuals 4 and 8. However, these could not explain the clinical phenotype because of either their functional and spatial properties or the fact that they were inherited from an unaffected parent. Additional information about these variants is presented in the Supplemental Note (also see Table S2).

According to the pLI (probability of loss-of-function intolerance) value in the ExAC Browser (version 0.3.1; accessed on September 10, 2017), *DPF2* seems to be highly intolerant to heterozygous loss-of-function variants (pLI = 1.00) and missense variants (Z score = 3.30).[35,36] Moreover, missense variants reported in the ExAC Browser are located mainly in the N-terminal domain of DPF2, whereas their frequency in PHD fingers is markedly lower (Figure 1A). Microdeletions encompassing *DPF2* have not been associated with a clinical phenotype to date.[37] In the Database of Genomic Variants (DGV),[38] two 149.5 and 242.1 kb deletions including *DPF2* have been described.

Figure 2. Facial Phenotype and Images from the Hands and Feet of Individuals with Variants in *DPF2*
(I1) Individual 1 at 9 years.
(I2) Individual 2 at 5 years, 5 months (image 1) and 16 years (images 2–5).
(I4) Individual 4 at 2 years (images 1, 4, and 5) and 12 years (images 2 and 3).
(I5) Individual 5 at 10 years, 5 months.
(I8) Individual 8 at 3 years, 9 months.
Note the hypoplasia of the fifth toenails (all individuals), hypoplasia of further toenails (I1, I4, and I8), hypoplasia of fingernails (I4 and I8), generalized brachydactyly (I4 and I8) or fifth-finger brachydactyly (I1 and I5), and clinodactyly (I2, I5, and I8). Facial dysmorphisms include sparse hair (I1, I4, and I8), a prominent forehead (I1, I4, and I8), hypertelorismus (I4), macrotia (I1, I2), prominent or low-set ears (I2, I4, and I8), a broad nose (I4 and I8), thick alae nasi (I1, I4, and I8), a thin upper lip (I2, I4, I5, and I8), a thick lower vermillion (I1 and I8), cleft lip palate (I8), and a broad and short philtrum (I1, I2, I4, and I8).

Table 1. Summary of the Clinical and Genetic Findings in Individuals with *De Novo DPF2* Variants

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 8
De novo DPF2 variant (GenBank: NM_006268.4)	c.827G>T (p.Cys276Phe)	c.990C>G (p.Cys330Trp)	c.1049G>A (p.Arg350His)	c.1037A>G (p.Asp346Gly)	c.1105T>C (p.Trp369Arg)	c.904+1G>T (p.?)	c.1099+1G>A (p.Asp340Glufs*12)	c.1066_1073del (p.Cys356Profs*5)
Exon or intron	exon 8	exon 9	exon 10	exon 10	exon 11	intron 8	intron 10	exon 10
Localization	PHD1	PHD2	PHD2	PHD2	PHD2	PHD1	PHD2	PHD2
Sex	male	female	female	male	male	female	female	male
Age at last clinical assessment	10 years	16 years	18 years, 6 months	NA	15 years	3 years, 2 months	7 years, 5 months	3 years, 9 months
Height	$<3^{rd}$ %	50^{th} % ($<3^{rd}$ % until puberty)	20^{th}–50^{th} %	$<3^{rd}$ %	$<0.4^{th}$ %	11^{th} %	15^{th} %	$<3^{rd}$ %
Weight	$<3^{rd}$ %	$<3^{rd}$ %	50^{th}–75^{th} %	$<3^{rd}$ %	$<0.4^{th}$ %	52^{th} %	39^{th} %	NA
OFC	50^{th} %	50^{th} %	–	$>97^{th}$ %	50^{th} %	$>97^{th}$ %	96^{th} %	50^{th} %
Brain anomalies	NA	+	NA	NA	NA	+	+	NA
Development								
Developmental delay	global	moderate global	moderate global	mild	global	mild to moderate	mild global	global
Cognition status	mild ID	moderate ID	moderate ID	borderline ID	moderate ID	precise estimation not possible (the individual is too young)	borderline ID	mild ID
Speech delay	+	+	+	+	+	+	+	+
Motor delay	–	+	+	+	–	–	+	+
Behavioral anomalies	+	–	+	–	+	–	–	–
Feeding problems	+	+	+	–	+	–	–	+
Muscular hypotonia	+	–	–	NA	–	+	+	+
Hearing loss	–	+	+	+	–	–	–	+
CSS-like facial features	+	–	+	+	–	coarse	coarse	+
Skeletal Anomalies								
Craniosynostosis	–	–	–	sagittal	sagittal	–	trigonocephaly (radiographic imaging was not performed)	–

(*Continued on next page*)

Table 1. *Continued*

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 8
Brachydactyly	only fifth finger	–	–	general	only fifth finger	only fifth finger	–	general
Clinodactyly	–	only fifth finger	–	–	only fifth finger	–	–	only fifth finger
Ectodermal Anomalies								
Sparse scalp hair	+	in childhood	+	+	–	NA	+	+
Nail hypoplasia or aplasia	forth and fifth toenails	right fifth toenail, small left fifth toenail	right fifth toenail, fourth and fifth fingernails, dysplasia of all nails	all toenails, fifth and index fingernails	fifth toenails, small thickened toenails	all toenails, fifth fingernails	fifth toenails	all toenails, fifth fingernails
Cardiac anomalies	–	+	–	+	–	+	+	–
Constipation	–	+	+	–	+	+	+	+
Recurrent otitis	+	–	+	NA	–	NA	+	+

Accordance with HGVS variant nomenclature was checked with Mutalyzer on September 10, 2017. The following abbreviations and symbols are used: +, present; –, absent; %, percentile; NA, not analyzed; OFC, occipitofrontal circumference; and PHD1 and PHD2, plant homeodomains.

The clinical phenotype of individuals 2–8 was similar to that of individual 1 and consistent with the phenotypic spectrum of CSS (Table 1 and Table S1). Global developmental delay was present in all individuals. Three (individuals 2, 3, and 5), one (individual 8), and two (individuals 4 and 7) individuals exhibited moderate, mild, and borderline intellectually disability, respectively. Individual 6 was 3 years and 2 months old at the last clinical evaluation, so a precise estimation of intellectual development was not possible. Speech delay was prominent among all affected children. In five (individuals 2–4, 7, and 8), motor milestones were delayed. Notably, hypoplasia of the fifth toenails, a central feature of CSS diagnosis, was observed in all individuals. Some individuals displayed additional skeletal and ectodermal CSS features, including hypoplasia of further toenails (individuals 4, 6, and 8) and/or fingernails (individuals 3, 4, 6, and 8), generalized or fifth-finger brachydactyly (individuals 4 and 8 or 5 and 6, respectively), and fifth-finger clinodactyly (individuals 2, 5, and 8) (see also Figure 2).

As is known about individuals with *ARID1B* variants, the CSS facial phenotype can be highly variable.[39] In our study, most individuals bearing *DPF2* variants presented with coarse facies. The most consistent facial features were sparse scalp hair (6/8), down-slanting palpebral fissures (6/8), thick or small alae nasi (6/8), a short or broad philtrum (6/8), large, prominent, low-set and/or posteriorly rotated ears (6/8), a prominent forehead (5/8), a broad nose (4/8), a wide mouth (4/8), a thin upper lip (4/8), a thick lower vermillion (4/8), and thick eyebrows (3/8) (Figure 2). Evaluation of growth parameters revealed that five of eight children presented with short stature, although height normalized during puberty in individual 2. In total, four individuals showed evidence of muscular hypotonia, and five manifested feeding problems. Behavioral anomalies—including stereotypic movements (individual 1); temper tantrums, obsessive-compulsive behavior, hyperactivity, poor sleep pattern, and stereotypic hand movements (individual 3); and fixations, temper tantrums, competitivity, and clinically suspected autism (individual 5)—varied across individuals. Furthermore, hearing impairment was described in four affected children. Notably, two individuals were diagnosed with sagittal craniosynostosis (individuals 4 and 5). Individual 7 was clinically diagnosed with trigonocephaly, but no radiographic imaging was performed to confirm a metopic craniosynostosis. To our knowledge, craniosynostosis is extremely rare in CSS, given that it was previously described in only two individuals with an *ARID1B* variant or a 2p25 deletion encompassing *SOX11*.[12,40] However, it is unclear whether it is a consistent manifestation in individuals bearing *DPF2* variants, given that both individuals 4 and 5 were selected from the DDD Study on the basis of this anomaly. Other common features included broad thumbs (3/8) and prominent fetal fingertip pads (3/8). Finally, mild to severe

constipation and recurrent otitis media were diagnosed in six and four of eight affected children, respectively. Three individuals underwent cerebral magnetic resonance imaging (MRI), which showed atrophy in the right cerebellar hemisphere (individual 2), a small pituitary gland (individual 6), and Arnold-Chiari malformation I (individual 7). No corpus callosum anomalies were noted (Table 1 and Table S1). Altogether, the clinical presentation of these individuals significantly overlaps that seen in CSS individuals. However, CSS was either clinically suspected or considered as a differential diagnosis only for individuals 1, 3, 7, and 8. The remaining individuals 2 and 4–6 were not clinical diagnosed with CSS, highlighting the value of trio exome sequencing as a clinical tool. Further phenotype-genotype correlation is not possible at the moment given the relatively small number of individuals.

In order to assess the effects of the *DPF2* missense variants on protein function and structure, we used a model based on the crystal structure of DPF2 and DPF3 bound to a histone peptide (Figure 1B) as a reference.[18,19] All five identified variants affect highly conserved residues within the C-terminal PHD1 and PHD2 motifs (PHD1: amino acids 272–327; PHD2: amino acids 329–374), which are responsible for the recognition of histone modifications. Strikingly, all affected amino acids are in very close proximity to a zinc binding site, most likely disrupting these sites and consequently the protein structure. This is especially evident for the affected tryptophan (Trp369) and both cysteine residues (Cys276 and Cys330), given that they directly disrupt a highly conserved residue of the C4HC3 motif, which anchors the two zinc atoms of the PHD fingers in a cross-brace topology (see also Figures 1C and 1D).[20,25] Interestingly, the affected Asp346 lies on the negatively charged PHD2 surface in the H3 peptide binding site and directly interacts with the H3 acetylated at lysine 14 (H3K14ac).[23] The remaining affected residues showed no direct interaction, but given the possible loss of the zinc binding sites and the altered conformation of PHD fingers, it can be speculated that the binding to H3 is weakened or even abolished (Figure 1B).

PHD fingers differentially recognize unmodified or methylated and acetylated lysine, as well as unmodified arginine residues in histone tails.[18,19,25,41] Proteins containing these motifs are implicated in the recruitment of chromatin remodellers and transcription factors.[25] The tandem PHD finger domain of DPF3, a DPF2 paralog, was shown to recruit the BAF chromatin-remodelling complex to regulate transcription by targeting N-terminal tails of both unmodified and modified acetylated and methylated lysines of histones H3 and H4. Variants disrupting the structure of one of the PHD fingers lead to the abolition of binding to unmodified H3 and H3 modifications. Yet, a single integral PHD is still able to recognize unmodified H4 or acetylated lysines of H4.[19,41] Similarly, intact DPF2 binds to unmodified and acetylated H4 and to unmodified, acetylated, and methylated H3, such that H3 acetylated at lysine 14 (H3K14ac) increases the affinity, whereas H3 methylated at lysine 4 (H3K4me3) reduces or inhibits the interaction.[19,23] It was recently reported that the tandem PHD finger of DPF2 is also responsible for the epigenetic regulation of transcription through the identification of rare histone modifications such as propionylation, butyrylation, and especially crotonylation, the last of which leads to a higher interaction with DPF2 than acetylation.[18] In a recent study, binding of an experimental construct bearing variants in both PHD fingers of DPF2 to H3 and H4 histone tails was abolished, whereas interaction with members of the BAF complex was not, suggesting that the major impact of these variants is the loss of histone binding.[23]

In order to experimentally investigate the impact of *DPF2* missense variants on histone interactions, we first generated a GST-tagged wild-type DPF2 fusion protein (GST-DPF2) and expression plasmids harboring three missense substitutions found in DPF2 individuals: p.Cys276Phe (GST-DPF2 C276F) in PHD1 and p.Cys330Trp (GST-DPF2 C330W) and p.Arg350His (GST-DPF2 R350H) in PHD2. The two remaining missense variants, p.Asp346Gly and p.Trp369Arg, were added at the latest stage of our study after completion of the functional analyses. To perform histone binding assays, we produced recombinant GST proteins by using a glutathione-S-transferase pull-down system. We assessed the interaction with the histones by using H3 and H4 biotinylated histone peptides, either unmodified or harboring specific modifications. We observed strong binding of the wild-type DPF2 to both unmodified and acetylated or methylated H3 (unmodified H3, H3K14ac, H3K9ac, H3K9/14ac, H3K9me2, and H3K9me3). Both H3 trimethylated at lysine 4 (H3K4me3) and H3 di- or trimethylated at lysine 27 (H3K27me2 or H3K27me3, respectively) showed either no or very weak interaction with DPF2. Unmodified and acetylated H4 at lysine 12 also showed binding, albeit weaker, to DPF2. Reciprocally, all three mutants abolished (H3K14ac and H3K9/14ac) or strongly attenuated (unmodified H3, H3K9ac, H3K9me2, and H3K9me3) the binding to H3, but not to unmodified or acetylated H4 (Figure 3A). These data suggest that the three missense variants examined indeed impair the structural integrity of either PHD1 or PHD2 and result in the aberrant reading of H3 histone modifications. Additionally, they confirm that the tandem PHD1-PHD2 finger works as an integral unit for the identification of H3 and H3 modifications by DPF2, as previously described for DPF3.[41] Considering that DPF2, like DPF3, is a subunit of the BAF chromatin-remodelling complex, we suggest that the identified missense substitutions in nucleosome-targeting modules disrupt the tandem PHD finger's functional cohesion and capacity to recognize H3 histone modifications and thus lead to misreading from the BAF complex and epigenetic deregulation of gene transcription, as shown for DPF3.[41]

Next, we determined whether the identified missense variants impair the sub-cellular localization of DPF2. To this end, we employed immunofluorescence confocal microscopy in HEK293 cells overexpressing GFP-tagged

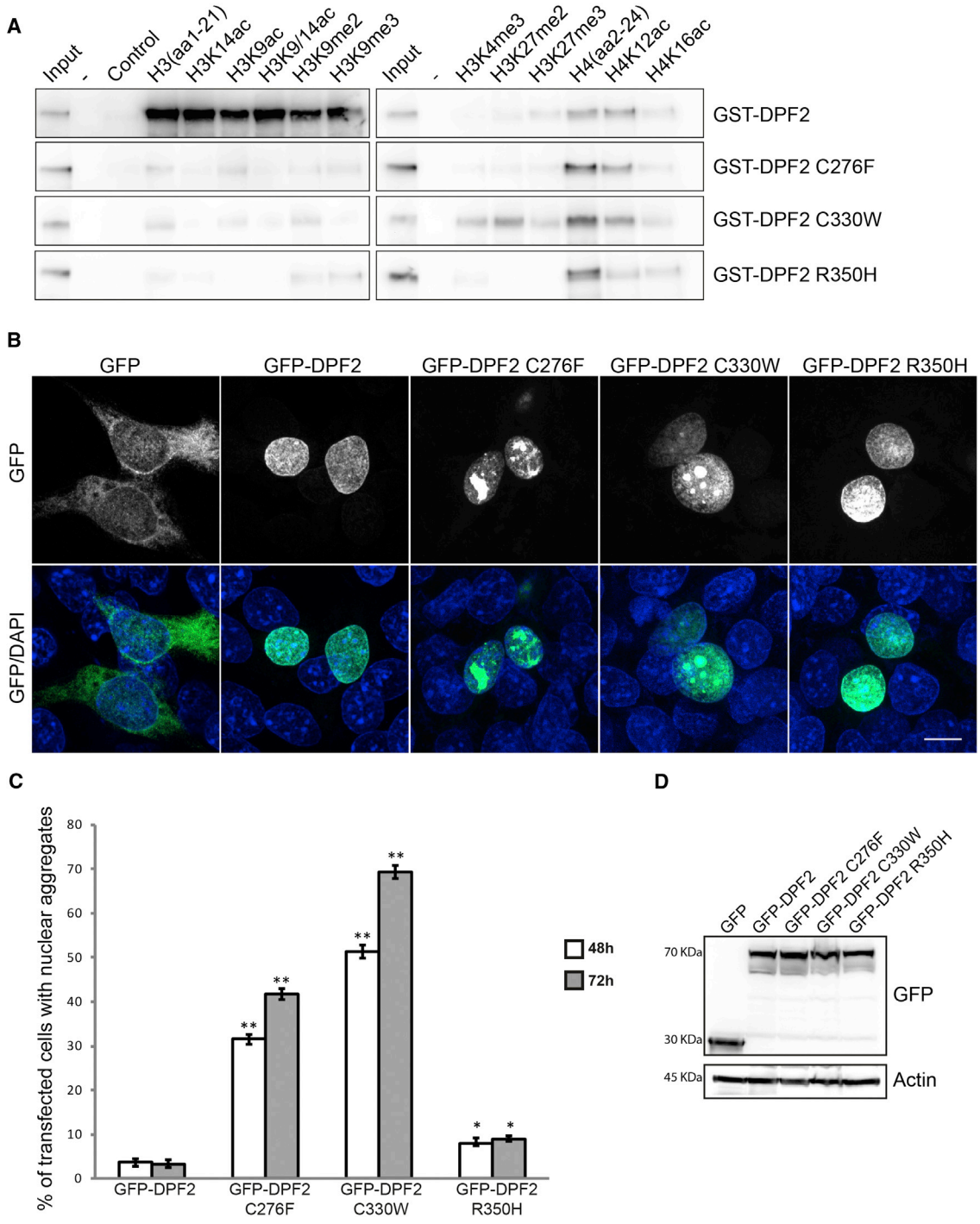

Figure 3. Functional Consequences of the *DPF2* Missense Variants on Histone Binding and Sub-nuclear Localization

(A) Western blot analysis of histone peptide pull-downs shows the absence or residual binding of the GST-DPF2 recombinant proteins harboring the three missense variants to unmodified, acetylated, or methylated H3 peptides (wild-type GST-DPF2 is shown for comparison). Note that the interaction with unmodified and acetylated H4 peptides was not affected. Purified GST-fusion proteins and histone peptides are indicated. Input represents 5% of the purified GST proteins used in the pull-downs. As a control, no peptide was added in the pull-down assay. GST-tagged wild-type DPF2 and mutants were generated by N-terminal subcloning from the pEGFP-C3 into a pGEX4T1(GST) vector via BglII and SalI restriction sites (Table S4). Expression of GST-DPF2 fusion proteins in BL21 bacteria was induced with 0.1 mM IPTG. Collected cells were resuspended in 25 mM HEPES (pH 7.5), 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, 10% glycerol, and proteinase inhibitor (Roche) and lysed with a French pressure cell press and then sonicated. Proteins were purified with Glutathione Sepharose 4B (GE Healthcare) according to the manufacturer's instructions and eluted with 10 mM glutathione (pH 8.5). The eluted protein was dialysed with the Slide-A-Lyzer MINI Dialysis Device (ThermoFisher Scientific) against 25 mM HEPES (pH 7.5), 500 mM KCl, 0.2 mM DTT, 1 mM EDTA, and 10% glycerol for 2 hr and the same buffer with 60% glycerol overnight. Histone-peptide binding assays were performed as described in Lange et al.[41] Biotinylated histone peptides were purchased from Millipore (unmodified H3, 12-403;

(legend continued on next page)

DPF2 (GFP-DPF2) or the PHD mutants. All transfected proteins were present in similar amounts (Figure 3D). Both wild-type and mutant proteins showed exclusively nuclear localization. This was diffuse for wild-type DPF2, whereas approximately 30% of cells expressing the missense variant in PHD1 (GFP-DPF2 C276F) and 50% of those with the missense variant in the second amino acid of PHD2 (EGFP-DPF2 C330W) formed distinct aggregate-like nuclear structures, which were negative for DAPI staining (Figures 3B and 3C). The percentage of cells exhibiting these abnormal nuclear accumulations peaked 72 hr after transfection (Figure 3C). The mutant harboring the alteration in the middle of the PHD2 finger (GFP-DPF2 R350H) had only a marginal effect on the formation of aggregates (Figures 3B and 3C). In order to exclude that the formed aggregates were technical artifacts relating to the GFP-tagged protein, we complementarily used a FLAG expression vector. The results were highly comparable (Figures S2A and S2B). The formation of aggregates was further validated in a second, COS7 cell line (Figure S2C). The observation of nuclear aggregates of DPF2 with a disrupted PHD finger structure is in accordance with a previous study that used HeLa cells to investigate the regulation of OCT4 by DPF2.[27] Protein misfolding can lead to protein aggregation, a major characteristic of several disorders, including neurodegenerative conditions.[43] We therefore speculate that the formation of nuclear aggregates in the two DPF2 mutants correlates with aberrant protein conformation, which could lead to protein dysfunction. Strikingly, in HEK293 cells, co-transfection experiments showed that the wild-type GFP-DPF2 was recruited to the aggregates formed by C276F and C330W FLAG-DPF2 mutants. This recruitement was accompanied by fewer cells showing aggregates than those co-expressing the FLAG-DPF2 mutants with control empty GFP (Figures S3A and S3B). Furthermore, co-expressed FLAG-BRG1 co-localized with the two GFP-DPF2 mutants within the aggregate-like structures (Figure S3C). The recruitment of wild-type DPF2 to the nuclear aggregates formed by mutant DPF2 proteins and the concomitant reduction of aggregate-exhibiting cells suggest an interaction between wild-type and mutant DPF2. This would be in line with a dominant-negative pathomechanism. In addition, the sequestration of the catalytic BRG1 subunit of the BAF complex in the aggregates suggests that these DPF2 variants could lead to the disruption of the physiological interactome of the BAF complex. Further experiments are necessary to elucidate whether the nuclear-aggregation phenotype is associated with the alterations that directly affect amino acids of the C4HC3 motif, given that the p.Arg350His variant, which has a minor effect, is the only one of the three studied substitutions to be located outside of it.

Variants in the two paralog BAF subunits of DPF2 (DPF1 and DPF3) have not been associated with disease in humans to date. Yet, alterations in genes encoding chromatin-related PHD-containing proteins, some of which even directly target the evolutionary conserved PHD motifs, have been implicated in many neurodevelopmental disorders. More specifically, mutations disrupting the PHD fingers encoded by *BRPF1* (MIM: 602410), *NSD1* (MIM: 606681), *ATRX* (MIM: 300032), *CREBBP* (MIM: 600140), *PHF6* (MIM: 300414), *KDM5C* (MIM: 314690), *KMT2A* (MIM: 159555), and *KMT2D* (MIM: 602113) cause syndromic intellectual disability (MIM: 617333), Sotos syndrome (MIM: 117550), alpha-thalassemia and mental retardation X-linked syndrome (MIM: 301040), Rubenstein-Taybi syndrome (MIM: 180849), Börjeson-Forssman-Lehmann syndrome (MIM: 301900), mental retardation, X-linked, syndromic, Claes-Jensen type (MIM: 300534), Wiedemann-Steiner syndrome (MIM: 605130), and Kabuki syndrome 1 (MIM: 147920), respectively.[44–49] Interestingly, some of the pathogenic variants found in other PHD-containing proteins affect homologous protein residues that overlap those altered in individuals with DPF2 variants (Figure S6 and Data S1). Dysregulated PHD proteins are also involved in additional human diseases (Figure 1D, Figures S4 and S5B, and Data S1).

In summary, the following lines of evidence support the pathogenicity of the identified *DPF2* variants: (1) their *de novo* occurrence, (2) the PHD finger mutational hotspot, (3) the intolerance of *DPF2* and especially its encoded tandem PHD finger domain to variation, and (4) the fact that functional analyses of the missense alterations showed abolished or attenuated H3 binding and the formation of

H3K9ac, 12-431; H3K14ac, 12-425; H3K9/14ac, 12-402; H3K9me2, 12-430; H3k9me3, 12-568; H3K4me3, 12-460, H3K27me2, 12-566; H3K27me3, 12-565; unmodified H4, 12-372) and Tebu-bio (H4K12ac, 12-0032; H4K16ac, 12-0033).

(B) Representative confocal immunofluorescence microscopy images of HEK293 cells overexpressing empty GFP vector (column 1) and wild-type DPF2 (column 2) and mutant (columns 3–5) GFP-tagged constructs. GFP was homogeneously distributed throughout the cell, whereas wild-type DPF2 and mutants exhibited exclusively nuclear localization. Note that C276F and C330W mutants led to the formation of nuclear protein aggregates, whereas R350H had a minor effect. The top panel shows GFP, and the lower panel shows merged pictures with DAPI. Scale bar, 10 μm. Wild-type DPF2 cDNA was amplified by PCR. The product was N-terminally inserted into the pEGFP-C3 vector with the In-Fusion HD Cloning Plus Kit (Clontech, Takara) according to the manufacturer's instructions. The variants in DPF2 were generated by mutagenesis PCR. Primers used for cloning and mutagenesis are described in Table S4. HEK293 cells were grown as described in Vasileiou et al.[42] 3×10^5 HEK293 cells were cultured on coverslips and transfected with 1 μg GFP and GFP-fusion plasmids with the use of polyethylenimine. After 48 hr, cells were fixed with 100% methanol at −20°C and stained with DAPI. Fluorescence images were acquired on an LSM 800 confocal laser scanning microscope (Carl Zeiss) with a 63× lens.

(C) HEK293 cells were examined 48 and 72 hr after transfection, and graphs indicate the percentage of cells exhibiting nuclear aggregates. Error bars represent the mean ± SEM, and significant values are indicated with *$p < 0.05$ or **$p < 0.01$ (unpaired Student's t test). 100 random cells were used per experiment ($n = 3$).

(D) Western blot analysis from lysates of HEK293 cells transfected with the indicated plasmids with the use of anti-GPF (top) and anti-panActin (bottom). Molecular weight (kDa) is indicated on the left of the western blotting panels.

nuclear DPF2 aggregates. Despite the intolerance to loss-of-function variants (as evidenced by the high pLI value), haploinsufficiency seems unlikely given that the annotation of deletions encompassing *DPF2* in normal populations indicates that a complete loss of one allele could be neutral. On the contrary, the clustering of all variants within an evolutionarily highly concerved region, as previously described for other BAF subunits (SMARCA4, SMARCA2, SMARCB1, and SMARCE1), the nuclear-aggregation phenotype, and the recruitment of wild-type DPF2 and BRG1 to the aggregates indicate a dominant-negative effect. The results of the expression analysis of the splice-site and frameshift *DPF2* variants found in affected individuals, suggesting that the truncated transcripts escape NMD, further support the dominant-negative hypothesis. Nevertheless, without protein analysis, we cannot completely exclude the possibility that a truncated RNA or protein instability leads to a specific loss of function of PHD fingers, at least for the truncating variants.

In conclusion, we have identified eight individuals who display a CSS phenotype and bear *de novo* variants in the DPF2 PHD finger domains. Our study further confirms the crucial role of PHD-finger-containing proteins in human neurodevelopmental disorders and strengthens the association between the etiology of CSS and variants affecting the function of subunits of the BAF chromatin-remodelling complex.

Accession Numbers

The accession numbers for the *DPF2* sequence variants reported in this article are LOVD: 00125791, 00131901, 00131902, 00131903, 00131904, 00131905, 00131906, and 00131907.

Supplemental Data

Supplemental Data include a Supplemental Note, Figures S1–S6, Tables S1–S4, and a data file and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.01.014.

Conflicts of Interest

M.T.C. is an employee of GeneDx.

Acknowledgments

We thank the individuals and their families for participating in this study. We also thank Juliane Hoyer and Michel Hadjihannas for useful advice. We are indebted to Daniela Schweitzer, Olga Zwenger, Angelika Diem, and Heike Friebel for excellent technical assistance. This work was supported by the German Federal Ministry of Research and Education (01GM1520A to A.R., 01GM1520B to N.B., 01GM1520D to T.S., and 01GM1520E to D.W.) as part of the Chromatin-Net Consortium, by Interdisciplinary Centre for Clinical Research Erlangen project E16 (A.R.), by the German Research Foundation (INST 410/91-1 FUGG to F.B.E), and by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre Programme (AOMW) and Wellcome Trust (Senior Investigator Award 102731 to the AOMW). The Deciphering Developmental Disorders Study presents independent research commissioned by the Wellcome Trust Sanger Institute (grant WT098051) and the HealthInnovation Challenge Fund (grant HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the Department of Health. The views expressed in this publication are those of the authors and not necessarily those of the Wellcome Trust or the Department of Health. The research team acknowledges the support of the NIHR through the Comprehensive Clinical Research Network.

Received: October 4, 2017
Accepted: January 17, 2018
Published: February 8, 2018

Web Resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
dbscSNV, https://sites.google.com/site/jpopgen/dbNSFP/
ExAC Browser, http://exac.broadinstitute.org/
GenBank, https://www.ncbi.nlm.nih.gov/genbank/
gnomAD browser, http://gnomad.broadinstitute.org/
Human Splicing Finder, http://www.umd.be/HSF3/
Leiden OpenVariation Database (LOVD), https://databases.lovd.nl/shared/genes/DPF2
Mutalyzer, https://mutalyzer.nl/
MutationTaster, http://www.mutationtaster.org/
NCBI, https://www.ncbi.nlm.nih.gov/
NNSPLICE version 0.9, https://omictools.com/nnsplice-tool
OMIM, http://www.omim.org/
PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/
RSCB Protein Data Bank, https://www.rcsb.org/pdb/home/home.do
SIFT, http://sift.jcvi.org/
SPIDEX, http://www.openbioinformatics.org/annovar/spidex_download_form.php

References

1. Kosho, T., Okamoto, N., Ohashi, H., Tsurusaki, Y., Imai, Y., Hibi-Ko, Y., Kawame, H., Homma, T., Tanabe, S., Kato, M., et al. (2013). Clinical correlations of mutations affecting six components of the SWI/SNF complex: detailed description of 21 patients and a review of the literature. Am. J. Med. Genet. A. *161A*, 1221–1237.
2. Coffin, G.S., and Siris, E. (1970). Mental retardation with absent fifth fingernail and terminal phalanx. Am. J. Dis. Child. *119*, 433–439.
3. Kosho, T., Miyake, N., and Carey, J.C. (2014). Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: historical review and recent advances using next generation sequencing. Am. J. Med. Genet. C. Semin. Med. Genet. *166C*, 241–251.
4. Hoyer, J., Ekici, A.B., Endele, S., Popp, B., Zweier, C., Wiesener, A., Wohlleber, E., Dufke, A., Rossier, E., Petsch, C., et al. (2012). Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am. J. Hum. Genet. *90*, 565–572.
5. Kosho, T., Okamoto, N.; and Coffin-Siris Syndrome International Collaborators (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1,

SMARCA4, SMARCE1, and ARID1A. Am. J. Med. Genet. C. Semin. Med. Genet. *166C*, 262–275.
6. Miyake, N., Tsurusaki, Y., and Matsumoto, N. (2014). Numerous BAF complex genes are mutated in Coffin-Siris syndrome. Am. J. Med. Genet. C. Semin. Med. Genet. *166C*, 257–261.
7. Santen, G.W., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., Kant, S.G., Snoeck, I.N., Peeters, E.A., Hilhorst-Hofstee, Y., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. Nat. Genet. *44*, 379–380.
8. Santen, G.W., Aten, E., Vulto-van Silfhout, A.T., Pottinger, C., van Bon, B.W., van Minderhout, I.J., Snowdowne, R., van der Lans, C.A., Boogaard, M., Linssen, M.M., et al.; Coffin-Siris consortium (2013). Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum. Mutat. *34*, 1519–1528.
9. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat. Genet. *44*, 376–378.
10. Tsurusaki, Y., Okamoto, N., Ohashi, H., Mizuno, S., Matsumoto, N., Makita, Y., Fukuda, M., Isidor, B., Perrier, J., Aggarwal, S., et al. (2014). Coffin-Siris syndrome is a SWI/SNF complex disorder. Clin. Genet. *85*, 548–554.
11. Wieczorek, D., Bögershausen, N., Beleggia, F., Steiner-Haldenstätt, S., Pohl, E., Li, Y., Milz, E., Martin, M., Thiele, H., Altmüller, J., et al. (2013). A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum. Mol. Genet. *22*, 5121–5135.
12. Hempel, A., Pagnamenta, A.T., Blyth, M., Mansour, S., McConnell, V., Kou, I., Ikegawa, S., Tsurusaki, Y., Matsumoto, N., Lo-Castro, A., et al.; DDD Collaboration (2016). Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. J. Med. Genet. *53*, 152–162.
13. Bramswig, N.C., Caluseriu, O., Lüdecke, H.J., Bolduc, F.V., Noel, N.C., Wieland, T., Surowy, H.M., Christen, H.J., Engels, H., Strom, T.M., and Wieczorek, D. (2017). Heterozygosity for ARID2 loss-of-function mutations in individuals with a Coffin-Siris syndrome-like phenotype. Hum. Genet. *136*, 297–305.
14. Chestkov, A.V., Baka, I.D., Kost, M.V., Georgiev, G.P., and Buchman, V.L. (1996). The d4 gene family in the human genome. Genomics *36*, 174–177.
15. Lessard, J., Wu, J.I., Ranish, J.A., Wan, M., Winslow, M.M., Staahl, B.T., Wu, H., Aebersold, R., Graef, I.A., and Crabtree, G.R. (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron *55*, 201–215.
16. Tando, T., Ishizaka, A., Watanabe, H., Ito, T., Iida, S., Haraguchi, T., Mizutani, T., Izumi, T., Isobe, T., Akiyama, T., et al. (2010). Requiem protein links RelB/p52 and the Brm-type SWI/SNF complex in a noncanonical NF-kappaB pathway. J. Biol. Chem. *285*, 21951–21960.
17. Wu, J.I. (2012). Diverse functions of ATP-dependent chromatin remodeling complexes in development and cancer. Acta Biochim. Biophys. Sin. (Shanghai) *44*, 54–69.
18. Xiong, X., Panchenko, T., Yang, S., Zhao, S., Yan, P., Zhang, W., Xie, W., Li, Y., Zhao, Y., Allis, C.D., and Li, H. (2016). Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2. Nat. Chem. Biol. *12*, 1111–1118.
19. Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature *466*, 258–262.
20. Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. Trends Biochem. Sci. *31*, 35–40.
21. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.
22. Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics *31*, 3997–3999.
23. Huber, F.M., Greenblatt, S.M., Davenport, A.M., Martinez, C., Xu, Y., Vu, L.P., Nimer, S.D., and Hoelz, A. (2017). Histone-binding of DPF2 mediates its repressive role in myeloid differentiation. Proc. Natl. Acad. Sci. USA *114*, 6016–6021.
24. Zhang, W., Xu, C., Bian, C., Tempel, W., Crombet, L., MacKenzie, F., Min, J., Liu, Z., and Qi, C. (2011). Crystal structure of the Cys2His2-type zinc finger domain of human DPF2. Biochem. Biophys. Res. Commun. *413*, 58–61.
25. Sanchez, R., and Zhou, M.M. (2011). The PHD finger: a versatile epigenome reader. Trends Biochem. Sci. *36*, 364–372.
26. Gabig, T.G., Mantel, P.L., Rosli, R., and Crean, C.D. (1994). Requiem: a novel zinc finger gene essential for apoptosis in myeloid cells. J. Biol. Chem. *269*, 29515–29519.
27. Liu, C., Zhang, D., Shen, Y., Tao, X., Liu, L., Zhong, Y., and Fang, S. (2015). DPF2 regulates OCT4 protein level and nuclear distribution. Biochim. Biophys. Acta *1853*, 3279–3293.
28. Deciphering Developmental Disorders, S.; and Deciphering Developmental Disorders Study (2017). Prevalence and architecture of de novo mutations in developmental disorders. Nature *542*, 433–438.
29. 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.
30. Bramswig, N.C., Lüdecke, H.J., Alanay, Y., Albrecht, B., Barthelmie, A., Boduroglu, K., Braunholz, D., Caliebe, A., Chrzanowska, K.H., Czeschik, J.C., et al. (2015). Exome sequencing unravels unexpected differential diagnoses in individuals with the tentative diagnosis of Coffin-Siris and Nicolaides-Baraitser syndromes. Hum. Genet. *134*, 553–568.
31. Gordon, C.T., Xue, S., Yigit, G., Filali, H., Chen, K., Rosin, N., Yoshiura, K.I., Oufadem, M., Beck, T.J., McGowan, R., et al. (2017). De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. Nat. Genet. *49*, 249–255.
32. Millan, F., Cho, M.T., Retterer, K., Monaghan, K.G., Bai, R., Vitazka, P., Everman, D.B., Smith, B., Angle, B., Roberts, V., et al. (2016). Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. Am. J. Med. Genet. A. *170*, 1791–1798.
33. Lindeboom, R.G., Supek, F., and Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat. Genet. *48*, 1112–1118.
34. Toma, K.G., Rebbapragada, I., Durand, S., and Lykke-Andersen, J. (2015). Identification of elements in human

long 3′ UTRs that inhibit nonsense-mediated decay. RNA *21*, 887–897.

35. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
36. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat. Genet. *46*, 944–950.
37. Nacinovich, R., Villa, N., Redaelli, S., Broggi, F., Bomba, M., Stoppa, P., Scatigno, A., Selicorni, A., Dalprà, L., and Neri, F. (2014). Interstitial 11q deletion: genomic characterization and neuropsychiatric follow up from early infancy to adolescence and literature review. BMC Res. Notes *7*, 248.
38. MacDonald, J.R., Ziman, R., Yuen, R.K., Feuk, L., and Scherer, S.W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. *42*, D986–D992.
39. Santen, G.W., Clayton-Smith, J.; and ARID1B-CSS consortium (2014). The ARID1B phenotype: what we have learned so far. Am. J. Med. Genet. C. Semin. Med. Genet. *166C*, 276–289.
40. Mignot, C., Moutard, M.L., Rastetter, A., Boutaud, L., Heide, S., Billette, T., Doummar, D., Garel, C., Afenjar, A., Jacquette, A., et al. (2016). ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. Brain *139*, e64.
41. Lange, M., Kaynak, B., Forster, U.B., Tönjes, M., Fischer, J.J., Grimm, C., Schlesinger, J., Just, S., Dunkel, I., Krueger, T., et al. (2008). Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev. *22*, 2370–2384.
42. Vasileiou, G., Ekici, A.B., Uebe, S., Zweier, C., Hoyer, J., Engels, H., Behrens, J., Reis, A., and Hadjihannas, M.V. (2015). Chromatin-Remodeling-Factor ARID1B Represses Wnt/β-Catenin Signaling. Am. J. Hum. Genet. *97*, 445–456.
43. Ross, C.A., and Poirier, M.A. (2004). Protein aggregation and neurodegenerative disease. Nat. Med. *10* (*Suppl*), S10–S17.
44. Gibbons, R.J., Bachoo, S., Picketts, D.J., Aftimos, S., Asenbauer, B., Bergoffen, J., Berry, S.A., Dahl, N., Fryer, A., Keppler, K., et al. (1997). Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. Nat. Genet. *17*, 146–148.
45. Jensen, L.R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A.R., Tariverdian, G., Chelly, J., Fryns, J.P., et al. (2005). Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am. J. Hum. Genet. *76*, 227–236.
46. Kalkhoven, E., Roelfsema, J.H., Teunissen, H., den Boer, A., Ariyurek, Y., Zantema, A., Breuning, M.H., Hennekam, R.C., and Peters, D.J. (2003). Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. Hum. Mol. Genet. *12*, 441–450.
47. Lower, K.M., Turner, G., Kerr, B.A., Mathews, K.D., Shaw, M.A., Gedeon, A.K., Schelley, S., Hoyme, H.E., White, S.M., Delatycki, M.B., et al. (2002). Mutations in PHF6 are associated with Börjeson-Forssman-Lehmann syndrome. Nat. Genet. *32*, 661–665.
48. Türkmen, S., Gillessen-Kaesbach, G., Meinecke, P., Albrecht, B., Neumann, L.M., Hesse, V., Palanduz, S., Balg, S., Majewski, F., Fuchs, S., et al. (2003). Mutations in NSD1 are responsible for Sotos syndrome, but are not a frequent finding in other overgrowth phenotypes. Eur. J. Hum. Genet. *11*, 858–865.
49. Yan, K., Rousseau, J., Littlejohn, R.O., Kiss, C., Lehman, A., Rosenfeld, J.A., Stumpel, C.T.R., Stegmann, A.P.A., Robak, L., Scaglia, F., et al.; DDD Study; and CAUSES Study (2017). Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation. Am. J. Hum. Genet. *100*, 91–104.

The American Journal of Human Genetics, Volume *102*

Supplemental Data

Mutations in the BAF-Complex Subunit *DPF2* Are Associated with Coffin-Siris Syndrome

Georgia Vasileiou, Silvia Vergarajauregui, Sabine Endele, Bernt Popp, Christian Büttner, Arif B. Ekici, Marion Gerard, Nuria C. Bramswig, Beate Albrecht, Jill Clayton-Smith, Jenny Morton, Susan Tomkins, Karen Low, Astrid Weber, Maren Wenzel, Janine Altmüller, Yun Li, Bernd Wollnik, George Hoganson, Maria-Renée Plona, Megan T. Cho, Deciphering Developmental Disorders Study, Christian T. Thiel, Hermann-Josef Lüdecke, Tim M. Strom, Eduardo Calpena, Andrew O.M. Wilkie, Dagmar Wieczorek, Felix B. Engel, and André Reis

Supplemental Note: Description of additional CNVs and SNVs

Microarray analysis in individual 1 showed a *de novo* 126 kb microdeletion in 14q21.3 (chr14:48,787,824-48,913,473) encompassing no known genes. The deletion was also previously identified in an individual of our control cohort, we thus considered it as a likely rare benign CNV. In individuals 2 and 3 were identified paternally inherited copy number aberrations. In individual 2 we detected a 328 kb microdeletion within 7q11.21 (chr7:63,449,550-63,777,276) encompassing no OMIM genes. In the absence of clinical features of the father we interpreted this as a rare benign CNV. Individual 3 had a microduplication of approximately 105 kb within 3p23-p22.3 (chr3:32,119,982-32,224,839) including the *GPD1L* gene (MIM: 611778). While missense variants in this gene have been correlated with Brugada-Syndrome 2 (MIM: 611777),[1] the affected child, as well as the father, did not present with any cardiac anomalies or arrhythmias. Furthermore, no cases of syncope or sudden unexplained death had been reported in the family, we thus also characterized it as non-causative. In individual 5 a missense variant c.3078G>C (p.(Lys1026Asn)) in *ASXL3* (MIM: 615115) gene was found, inherited from the healthy father. This substitution has been observed in ExAC and gnomAD once (rs372219878) and was predicted *in silico* as potentially affecting function. *De novo* mutations in *ASXL3* gene have been described in patients with a novel syndrome with phenotypic similarities with Bohring-Opitz syndrome, characterized by severe psychomotor delay, feeding difficulties, small size at birth and microcephaly.[2] Recently the intellectual disability syndrome due to *ASXL3* mutations was recognized as a distinct clinical phenotype, called Bainbridge-Ropers syndrome (MIM: 615485).[3] Nevertheless, all pathogenic mutations reported until recently are truncating and are located within two evolutionarily conserved regions with phosphorylation sites and a serine-rich motif, located upstream of the identified substitution in individual 5.[2] Taking the aforementioned data together, the missense alteration c.3078G>C (p.(Lys1026Asn)) in *ASXL3* seems to be a rare variant not affecting function and we thus did not consider it as relevant to the phenotype.

Individual 6 further carried a maternally inherited *GDF1* (MIM: 602880) variant c.681C>A (p.(Cys227*)) (rs121434422). This nonsense variant has been described as pathogenic four times in ClinVar database (Variation ID: 6747) in patients with dextro-looped transposition of the great arteries with autosomal dominant inheritance (MIM: 613854), as well as in individuals with autosomal recessive right atrial isomerism (MIM: 208530), a heterotaxy syndrome, including asplenia.[4,5] Although the mother was healthy without any cardiac anomalies and the bicuspid aortic valve as well as the dilated ascending aorta diagnosed in

individual 6 are not the major cardiac manifestations of patients with this mutation, we cannot exclude a possible contribution to the described cardiac anomalies. Yet the remaining clinical features cannot be explained by this aberration. Moreover, microarray analysis revealed a *de novo* duplication of 652 kb (chrX:92,416,787-93,069,211) in Xq21.32, encompassing two genes; *FAM133A* without OMIM phenotype and *NAP1L3* (MIM: 300117), the latter gene being homologous with the genes of nucleosome assembly protein family (NAPs) encoding for brain specific proteins. *NAP1L3* is closely linked to a region (Xq21.3-q22) of genes responsible for X-linked intellectual disability, but to our knowledge neither *NAP1L3* nor other members of NAP family have been implicated in developmental-delay disorders.[6] We characterized this as a variant of unknown significance (VUS), but given that the aberration is a gain, we consider the variant unlikely to have clinical significance. In individual 7, a *de novo*, heterozygous missense variant c.4885C>T (p.(Arg1629Cys)) in the chromatin modifier *SETD1B* (MIM: 611055), a subunit of histone methylotransferase complex, was identified. Microdeletions in 12q24.31 encompassing *SETD1B* have been associated with syndromic intellectual disability. The major clinical phenotype of the patients reported includes moderate to severe intellectual disability, autism, epilepsy, behavioural problems and muscular hypotonia. Labonne et al. reported the smallest 360 kb microdeletion in this region including 6 candidate genes and proposed *SETD1B* as one of the two most compatible candidates to the clinical phenotype.[7,8] *SETD1B* seems not to be a conserved gene (ExAC pLI=0.02 and Z score=1.57), making a pathogenic relevance of the variant c.4885C>T (p.(Arg1629Cys)) unlikely. Also, variants in *SETD1B* or in the other candidate genes in this region have not been associated with disease in humans to date; consequently the underlying mechanism of deletions in this region behind the clinical phenotype remains elusive. Finally, individual 7 presented with borderline ID without evidence of seizures or autistic behaviour which is not consistent with the phenotype described in patients with 12q24.31 microdeletions (Table S2).

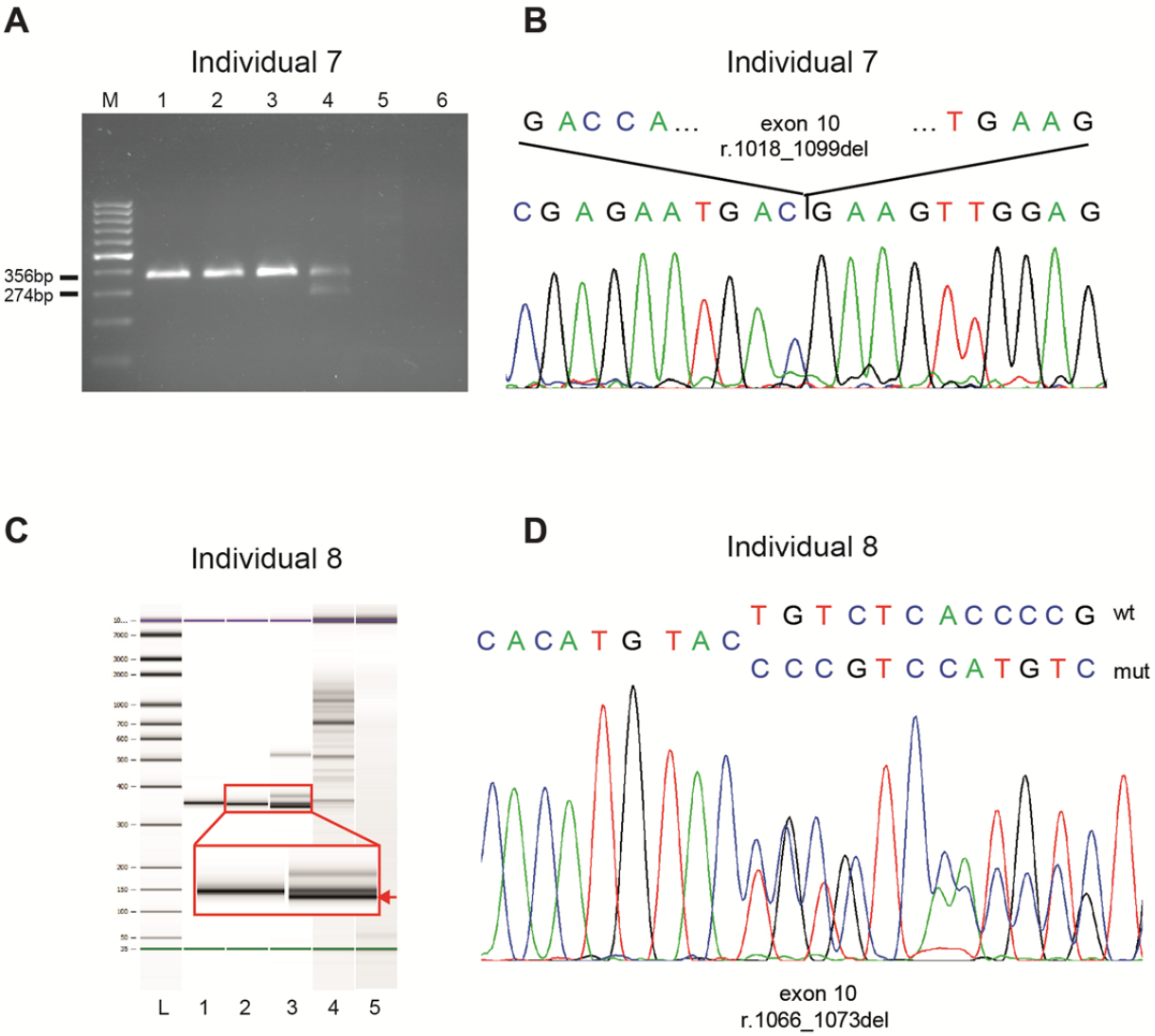

Figure S1. RT-PCR analysis of *DPF2* in individuals 7 and 8.

(A) Exon spanning RT-PCR on RNA from PAXgene stabilized lymphocytes of individual 7 using primers in exon 9 (DPF2_Exon9_RT_F) and 11 (DPF2_Exon11_RT_R) (Table S4) resulted in an additional aberrant product of 274 bp in individual 7 (lane 4), whereas controls (lanes 1–3) showed only the expected 356 bp fragment (lane 5, genomic DNA control; lane 6, no template control; M, size standard 100 bp).

(B) Sequencing electropherogram of RT-PCR amplified aberrant product verifies the skipping of exon 10 (r.1018_1099del).

(C) Agilent Bionalyzer electrophoresis analysis of exon spanning RT-PCR on RNA from PAXgene stabilized lymphocytes of individual 8 using primers described in Figure S1A. Note the presence of the additional aberrant transcript of 348 bp in individual 8 (lane 3). Controls (lanes 1–2) showed only the expected 356 bp transcript. Lane 4, genomic DNA control; lane 5, no template control; L, size standard). Enlargement of the two fragments (red outline); the aberrant product is indicated with the red arrow.

(D) Sequencing electropherogram of RT-PCR product from the affected Individual 8 confirms the 8 bp deletion (r.1066_1073del).

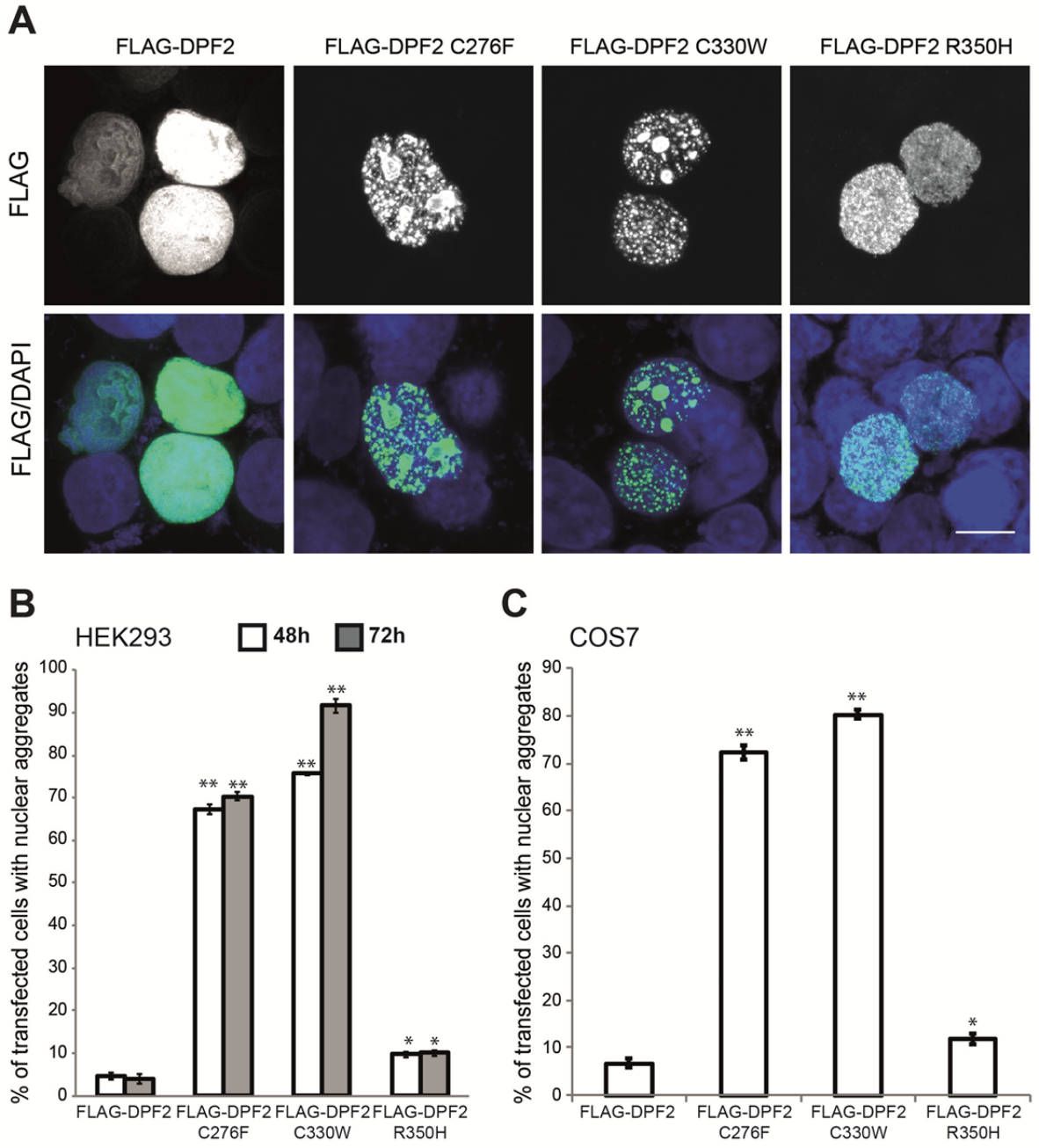

Figure S2. Validation of aggregate formation in cells transfected with FLAG-tagged DPF2 mutants.

(A) Representative confocal immunofluorescence microscopy images of HEK293 cells expressing FLAG-tagged DPF2 and mutants. Note the formation of nuclear aggregate-like structures in cells transfected with the indicated mutants (columns 2-4) as compared to the wild type DPF2 (column 1). Top panel shows staining with antibody against FLAG and lower panel merged pictures with DAPI. Scale bar 10 µm. FLAG-tagged proteins were generated by subcloning from the pGEX4T1 vector into a pCMV-Tag2b (FLAG) vector via EcoRI/SalI restriction sites. Primers used for cloning and mutagenesis are described in Table S4. Transfection with FLAG-tagged DPF2 plasmids and immunofluorescence were performed as described in Figure 3B, but using Lipofectamine (Invitrogen). Cells were incubated with mouse monoclonal anti-FLAG M2 primary antibody (Sigma-Aldrich) followed by incubation with goat anti-mouse conjugated to Alexa Fluor 488 (Life Technologies).

(B) HEK293 cells were examined after 48 and 72 hours post transfection and graphs indicate the percentage of cells exhibiting nuclear aggregates. Error bars represent mean ± S.E.M. and significant values are indicated with *p<0.05, **p<0.01 by unpaired student´s t-test. 100 random cells per experiment, n=3.

(C) Quantification of COS7 cells after 48 hours post transfection with indicated plasmids showing the percentage of cells exhibiting nuclear aggregates. 100 random cells per experiment, n=3.

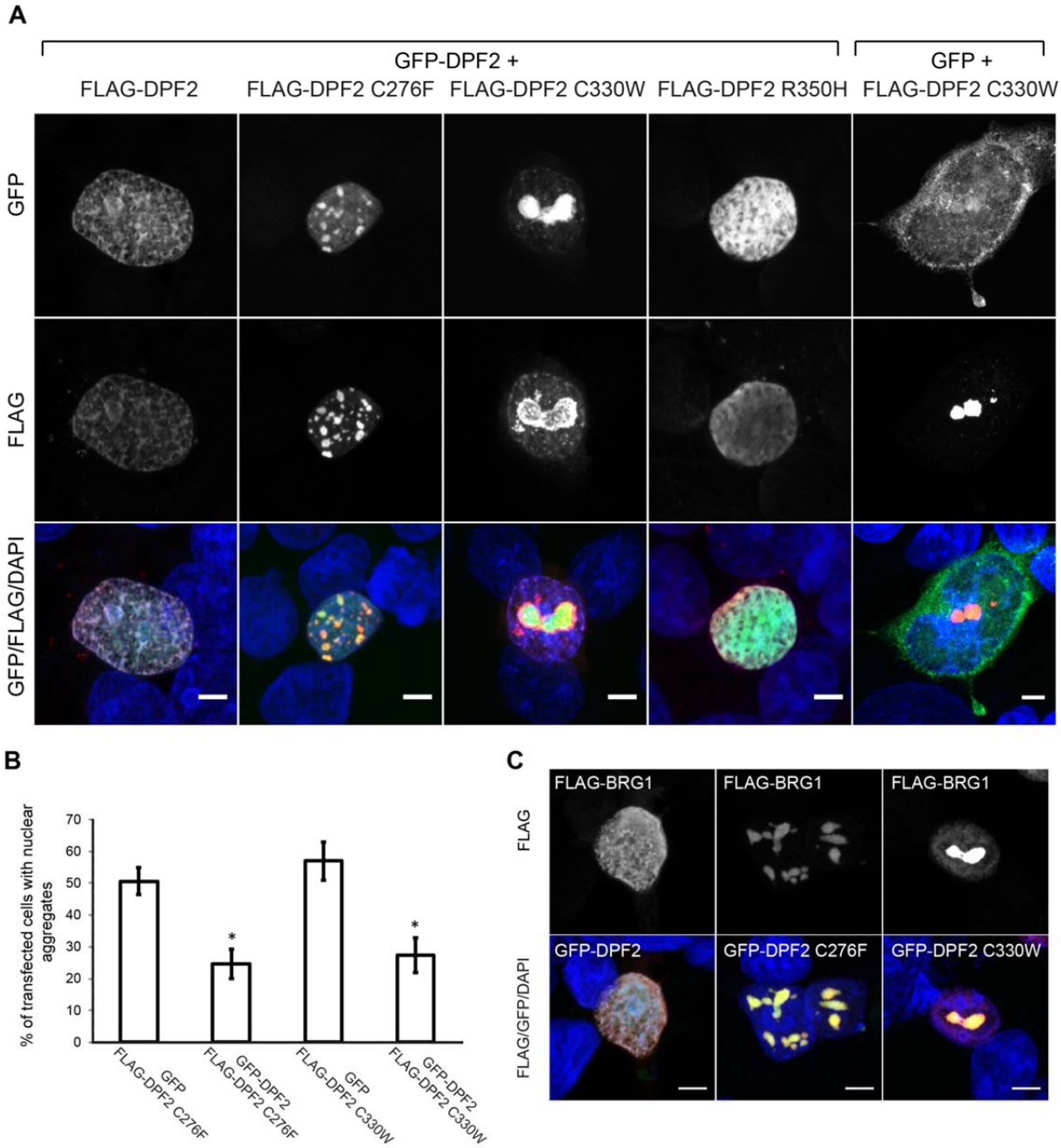

Figure S3. C276F and C330W mutants recruit wild type DPF2 and BRG1 to the aggregates.

(A) Representative confocal immunofluorescence microscopy images of HEK293 cells co-transfected with GFP-tagged DPF2 (columns 1-4) or empty GFP vector (column 5) together with FLAG-tagged DPF2 and mutants (columns 2-5). Note the recruitment of wild type GFP-DPF2 to the nuclear aggregate-like structures in cells co-transfected with FLAG-tagged mutants (columns 2 and 3). Empty GFP vector is not sequestered to the aggregates (column 5). Top panel shows GFP, middle panel staining with antibody against FLAG (FLAG M2 primary antibody; Sigma-Aldrich) and lower panel merged pictures with GFP/FLAG/DAPI. Scale bar 5 µm.

(B) Quantification of experiments conducted in (A) after 48 hours post transfection. Error bars represent mean ± S.E.M. and significant values are indicated with *$p<0.05$ by unpaired student's t-test. 100 random cells per experiment, n=3.

(C) Representative confocal immunofluorescence microscopy images of HEK293 cells co-expressing GFP-tagged DPF2 (column 1) and mutants (columns 2-3) together with FLAG-tagged BRG1 (Addgene).[9] Note the co-localization of BRG1 with the indicated mutants in the nuclear aggregates (columns 2-3). Top panel shows staining with antibody against FLAG and lower panel merged pictures with FLAG/GFP/DAPI. Scale bar 5 μm.

Cladogram PHD finger domains

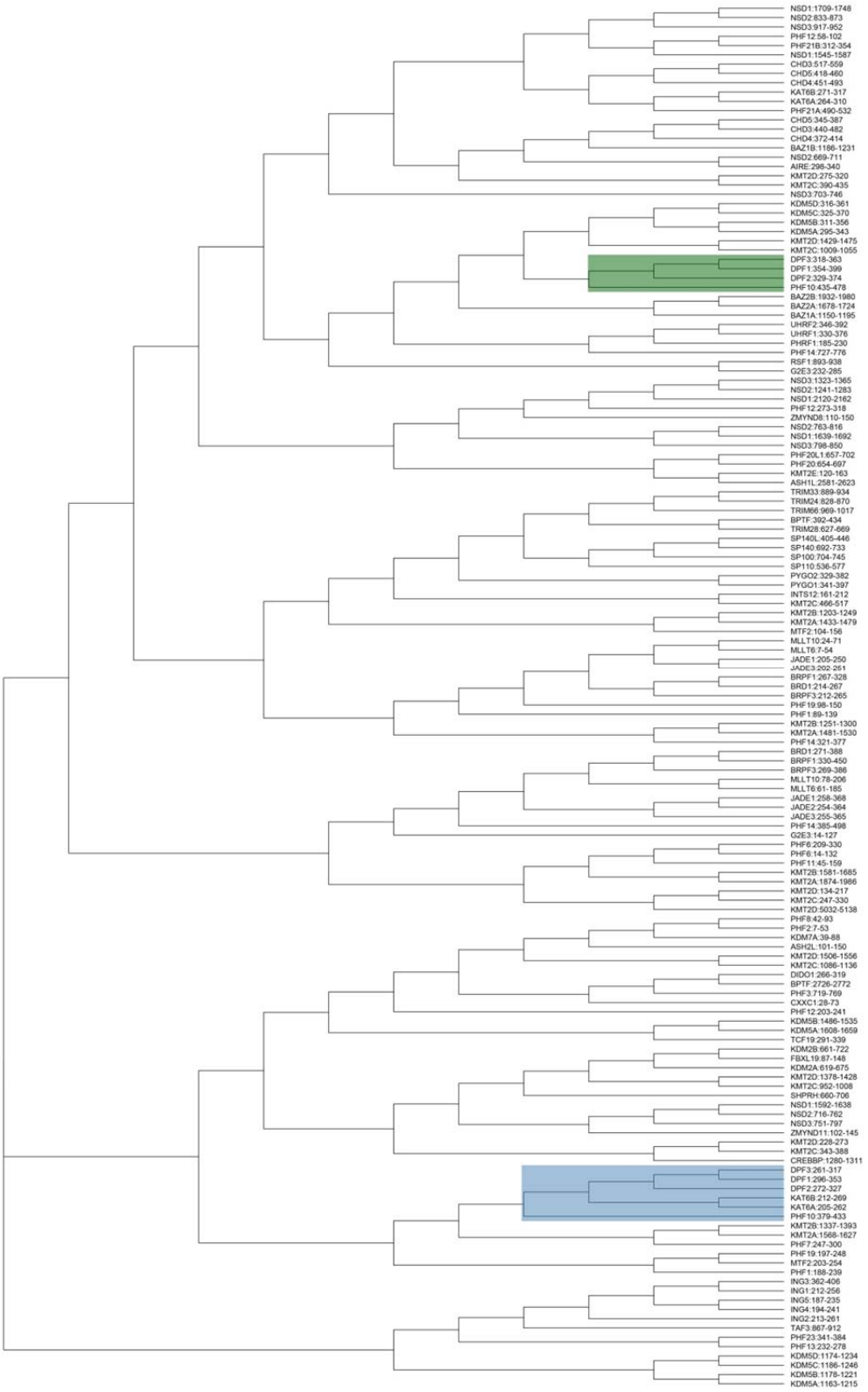

Figure S4. Cladogram of PHD finger domains used to find DPF2 paralog proteins.

To find putative human paralog proteins of DPF2 with similar tandem PHD fingers for the multiple amino-acid sequence alignment we compiled a list of known proteins with PHD finger domains (see also Data S1, sheet ´PHD_finger_proteins_(PHF)´) using different sources (see also Data S1, sheet ´data´) and extracted the amino acid sequence of their PHD finger domains based on the NCBI annotations using bedtools v2.26.0.[10] These sequences were then aligned using ClustalW[11] the msa package in R.[12] The cladogram was plotted using the ape package [13] and ggtree package in R.[14] The node for the first tandem PHD finger was marked in blue and the node for the second PHD finger was marked in green. Based on these results and the visual inspection of the protein domain structures the proteins DPF1 (NP_001128627.1), DPF3 (NP_001267471.1), KAT6A (NP_006757.2), KAT6B (NP_036462.2) and PHF10 (NP_060758.2) were selected as putative paralogs of DPF2 (NP_006259.1) and used for multiple sequence alignment (Figure 1D and Figure S5B).

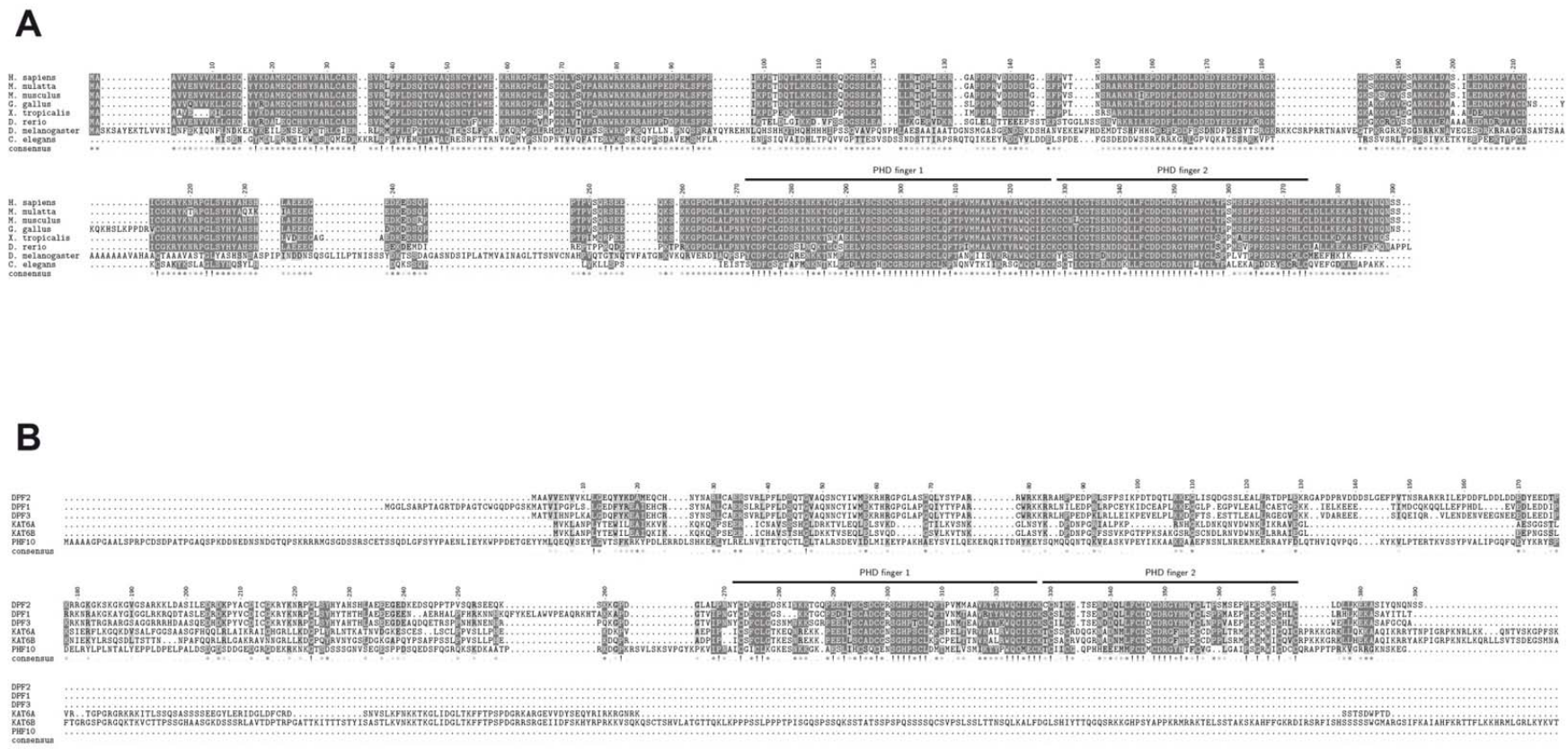

Figure S5. Full multiple sequence alignment of DPF2 orthologs and paralogs.

(A) Full multiple sequence alignment of DPF2 orthologs. Protein sequences were obtained from the NCBI (see Data S1, sheet ´DPF2_orthologs´) and ClustalW[11] from the msa package[12] in R was used for alignment. **(B)** Full multiple amino-acid sequence alignment of DPF2 and putative human paralog proteins with similar tandem PHD fingers. Protein sequences were obtained and alignment was performed as described above (see Data S1, sheet ´PHD_finger_proteins_(PHF)´)

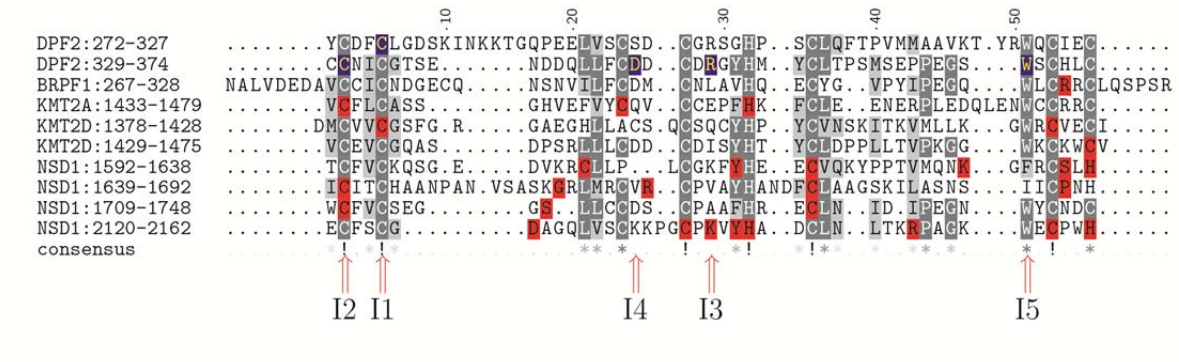

Figure S6. Multiple sequence alignment of PHD finger domains from intellectual disability associated genes with (likely) pathogenic missense variants from ClinVar

Protein sequences were obtained from the NCBI while missense variants of class 4 or 5 in genes associated with an intellectual disability phenotype in OMIM containing a PHD finger domain were obtained from ClinVar on 2017-11-26 (see Data S1, sheet ´PHD_finger_proteins_(PHF)´). Only missense variants in canonic PHD fingers (PHD finger 1-6 in sheet ´PHD_finger_proteins_(PHF)´) were considered. ClustalW[11] from the msa package in R[12] was used for alignment. The positions with de novo missense variants in *DPF2* in the 5 individuals (I1-I5) are indicated with a red arrow. The homologous amino acid position affected by the variant c.827G>T (p.(Cys276Phe)) in the first PHD finger (272..327) of DPF2 (I1) is also affected by the c.4148G>A (p.(Cys1383Tyr)) variant (ClinVar Variation ID: 419661) in the PHD finger 3 (1378..1428) of KMT2D. The homologous amino acid position affected by the variant c.990C>G (p.(Cys330Trp)) in the second PHD finger (329..374) of DPF2 (I2) is also affected by the c.4301G>C (p.(Cys1434Ser)) variant (ClinVar Variation ID: 429266) in the PHD finger 1 (1433..1479) of KMT2A and by the c.4918T>A (p.(Cys1640Ser)) and c.5129G>T (p.(Cys1710Phe)) variants (ClinVar Variation IDs: 235776, 373004) in the PHD finger 3 and PHD finger 4 (1639..1692; 1709..1748) of NSD1. The homologous amino acid position affected by the variant c.1049G>A (p.(Arg350His)) in the second PHD finger (329..374) of DPF2 (I3) is also affected by the c.6418A>G (p.(Lys2140Glu)) variant (ClinVar Variation ID: 159418) in the PHD finger 5 (2120..2162) of NSD1. Additionally other variants in these genes affect highly conserved amino acid positions in the Cys4-His-Cys3 (C4HC3) architecture.

Table S1. Detailed description of clinical and genetic findings in individuals with *de novo DPF2* variants

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 8
***DPF2* variant (NM_006268.4)**	c.827G>T p.(Cys276Phe)	c.990C>G p.(Cys330Trp)	c.1049G>A p.(Arg350His)	c.1037A>G p.(Asp346Gly)	c.1105T>C p.(Trp369Arg)	c.904+1G>T p.?	c.1099+1G>A p.Asp340Glufs*12	c.1066_1073del p.Cys356Profs*5
Exon	8	9	10	10	11	intron 8	intron 10	10
Genomic position (hg19)	chr11:g.65113452G>T	chr11:g.65113803C>G	chr11:g.65116352G>A	chr11:g.65116340A>G	chr11:g.65119159T>C	chr11:g.65113530G>T	chr11:g.65116403G>A	chr11:g.65116369_65116376del
Localization	PHD1	PHD2	PHD2	PHD2	PHD2	PHD1	PHD2	PHD2
Inheritance	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*
Sex	male	female	female	male	male	female	female	male
Family history	negative	negative	negative	negative	negative	negative	negative	younger maternal half-brother with *de novo* microdeletion 22q11
Abnormalities during pregnancy	-	-	-	-	-	mild polyhydramnios	oligohydramnios, singular umbilical artery, elevated nuchal translucency	-
Birth parameters (length, weight, OFC)	18th ct, 12th ct, 25th ct	16th ct, 9th ct, 29th ct	NA, NA, NA	NA, NA, 47th ct	NA, 92th ct, NA	3rd ct, 13th ct, large (precise size not known)	11th ct, 9th ct, 7th ct	10-25th ct, 10th ct, 50th ct
Age at last clinical assessment	10 years	16 years	18 years, 6 months	NA	15 years	3 years, 2 months	7 years, 5 months	3 years, 9 months
Height	< 3rd ct	50th ct (< 3rd ct until puberty)	20-50th ct	< 3rd ct	< 0,4th ct	11th ct	15th ct	< 3 ct
Weight	< 3rd ct	< 3rd ct	50-75th ct	< 3rd ct	< 0,4th ct	52th ct	39th ct	NA
OFC	50th ct	50th ct	-	>97th ct	50th ct	>97th ct	96th ct	50th ct
Brain MRI scan	NA	right cerebellar hemisphere atrophy	NA	NA	NA	small pituitary gland	Arnold-Chiari malformation I	NA
Development								
Developmental delay	global	moderate global	moderate global	mild	global	mild/moderate	mild global	global
Cognition status	mild ID	moderate ID	moderate ID	borderline ID	moderate ID	precise estimation not possible (the patient is too young)	borderline ID	mild ID
First words	48 months	36 months	48 months	36 months	24 months	~24 months	33 months	18 months (mama)- no other words for many months
First sentences	84 months	53 months	60 months	NA	NA	not speaking in sentences yet	NA	36 months
Age of walking	17 months	24 months	23 months	22 months	18 months	17 months	23 months	21 months
Behavioral anomalies	stereotypes	-	temper tantrums, hyperactivity, obsessive compulsive behavior. poor sleep pattern, stereotypic movements of hands	-	fixations, temper tantrums, competitive behavior, autism	-	-	-
Feeding problems	+	+	+	-	+	-	-	+
Beginning of feeding problems	at 6 months	infancy	NA		infancy			infancy (due to bilateral cleft lip palate)

Duration of feeding problems	1-2 years	ongoing due to gastro-esophageal reflux	ongoing		until solid food introduced			ongoing
Muscular hypotonia	+	-	-	NA	-	+	+	+
Hearing loss	-	+ (high frequency)	+ (congenital,severe left-sided)	+ (mild conductive)	-	-	-	+
Vision problems	hypermetropia, photophobia	unilateral intermittent strabismus	-	-	strabismus, hypermetropia	-	myopia	-
Craniofacial anomalies								
Prominent forehead	+	-	-	+	-	+	+	+
Thick eyebrows	+	-	+	+	-	NA	-	-
Long eyelashes	+	-	+	-	-	NA	+	+
Hypertelorism	-	-	-	+	-	-	+	-
Epicanthic fold	-	-	-	-	-	NA	+	+
Down-slanting palpebral fissures	+	+	+	+	+	NA	-	+
Flat nasal bringe	-	-	-	-	-	+	+	+
Broad nose	-	-	+	+	beaked nose	-	+	+
Upturned nasal tip	-	-	-	+	-	-	-	-
Thick alae nasi	+	-	+	+	small alae nasi	-	+	+
Broad philtrum	-	-	+	+	-	-	-	+
Short philtrum	+	+	-	-	-	+	-	+
Thin upper lip	-	+	-	+	+	-	-	+
Thin upper vermillion	-	-	-	+	-	-	-	-
Thick lower vermillion	+	-	+	-	-	-	+	+
Wide mouth	+	-	-	+	-	-	+	+
Downturned mouth	-	-	-	-	-	-	+	-
Macrotia	+	+	+	-	-	-	-	-
Other ear anomalies	-	low-set and/or posteriorly rotated ears	stenosis of external auditory canal	low-set and/or posteriorly rotated ears	prominent, cup-shaped, low-set ears and abnormality of the pinna	creased earlobes with overfolded helices	-	prominent, low-set and/or posteriorly rotated ears
Skeletal anomalies								
Craniosynostosis	-	-	-	sagittal	sagittal	-	trigonocephaly (radiographic imaging was not performed)	-
Delayed bone age	NA	NA	NA	NA	-	+	+	NA
Scoliosis	-	-	-	+	-	-	-	-
Pectus excavatum	-	+	-	-	+	-	-	-
Brachydactyly	only of 5th finger	-	-	general	only of 5th finger	only of 5th finger	-	general

Clinodactyly	-	only of 5th finger	-	-	only of 5th finger	-	-	only of 5th finger
Complete absent of 5th distal phalanx	-	-	-	possibly	-	-	-	-
Joint laxity	+	-	-	+	-	NA	NA	+
Ectodermal anomalies								
Sparse scalp hair	+	in childhood	relatively	relatively	-	NA	+ (and fine hair)	+
Hypertrichosis	-	-	+ (upper lip, arms, legs and back)	-	-	-	-	-
Nail hypoplasia/Aplasia	4th and 5th toenails	right 5th toenail, small left 5th toenail	right 5th toenail, 4th and 5th fingernails, dysplasia of all nails	all toenails, 5th and index fingernails	5th toenails, small thickened toenails	all toenails, 5th fingernails	5th toenails	all toenails, 5th fingernails
delayed primary dentition	NA	+	NA	-	-	NA	+	-
delayed permanent dentition	NA	-	NA	+	-	NA	+	+
Microdontia	+	-	-	-	-	NA	-	+
C**ongenital anomalies**								
Intestinal anomalies	-	gastroesophageal reflux (gastrostomy), constipation	severe constipation since infancy	-	severe constipation (colostomy at 3 years and colectomy with ileostomy at 13 years)	constipation	constipation	constipation
hernia	-	-	umbilical	inguinal	-	-	-	-
Cardiac anomalies	-	mild pulmonary stenosis, PFO	-	VSD	-	bicuspid aortic valve, dilated ascending aorta	aortic valve dysplasia	-
Renal anomalies	-	-	-	large kidneys	-	-	-	-
Laryngo/Tracheomalacia	-	laryngomalacia		-	-	-	laryngomalacia	-
Dermatological anomalies	Cutis marmorata	Cutis marmorata	-	multiple melanocytic naevi, dry skin palms of hands, supernumerary nipple, "bronzed" skin	capillary	capillary hemangioma on forehead, orbital creases	-	neurodermatitis
Recurrent otitis	+ (otitis media)	-	+	NA	-	NA	+ (otitis media)	+(otitis media)
Other	-	low hanging columella, bilateral single palmar creases	prominent fetal fingertip pads, low anterior and posterior hairlines, synophrys	very thick gums, broad thumbs	prognathism, broad thumbs, bilateral undescended testes IGF-1 deficiency	prominent fetal fingertip pads, short broad thumbs, pes planus, decrease of calf muscle mass	-	low anterior hairline, bilateral cleft lip palate, prominent fetal fingertip pads, widely spaced nipples, undescended testes, Perthes disease, iron deficiency,

Table S2. Additional CNVs and gene alterations identified in DPF2 individuals

	CNVs	Gene variants
Individual 1	arr[hg19] 14q21.3(48,787,824-48,913,473)x1	-
Individual 2	arr[hg19]7q11.21(63,449,550-63,777,276)x1	-
Individual 3	arr[hg19] 3p23p22.3(32,119,982-32,224,839)x3	-
Individual 4	-	-
Individual 5	-	NM_030632.2 (*ASXL3*): c.3078G>C (p.(Lys1026Asn))
Individual 6	arr[hg19] Xq21.32(92,416,787-93,069,211)x3	NM_001492.5 (*GDF1*): c.681C>A (p.(Cys227*))
Individual 7	-	NM_015048.1(*SETD1B*): c.4885C>T (p.(Arg1629Cys))
Individual 8	-	-

Table S3: Computational analysis of DPF2 variants

ID	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 8
CHROM	chr11:g.65113452 G>T	chr11:g.65113803 C>G	chr11:g.65116352 G>A	chr11:g.65116340 A>G	chr11:g.65119159 T>C	chr11:g.65113530 G>T	chr11:g.65116403 G>A	chr11:g.65116369_65116376del
GENE	*DPF2*	*DPF2*	*DPF2*	*DPF2*	*DPF2*	*DPF2*	*DPF2*	*DPF2*
FEATUREID	NM_006268.4	NM_006268.4	NM_006268.4	NM_006268.4	NM_006268.4	NM_006268.4	NM_006268.4	NM_006268.4
EFFECT	missense variant	missense_variant	missense_variant	missense_variant	missense_variant	splice_donor_variant&intron_variant	splice_donor_variant&intron_variant	frameshift_variant
HGVS_C	c.827G>T	c.990C>G	c.1049G>A	c.1037A>G	c.1105T>C	c.904+1G>T	c.1099+1G>A	c.1066_1073del
HGVS_P	p.(Cys276Phe)	p.(Cys330Trp)	p.(Arg350His)	p.(Asp346Gly)	p.(Trp369Arg)	p.?	p.Asp340Glufs*12	p.Cys356Profs*5
gnomadr201_AC	na	na	na	na	na	na	na	na
Cosmicv81_CNT	na	na	1	na	na	na	na	na
CADDv1.3_phred	31.0	27.5	35.0	34.0	27.2	26.1	28.2	35.0
M_CAP_score	0.421	0.956	0.169	0.456	0.602	na	na	na
REVEL_score	0.525	0.876	0.736	0.853	0.83	na	na	na
SIFT_score	0.378	0.0	0.0	0.004	0.0	na	na	na
Polyphen2_HVAR_score	0.0	0.999	0.991	0.989	0.992	na	na	na
MutationTaster_score	1.0	1.0	1.0	1.0	0.999	1.0	1.0	na
spidex_dpsi_zscore	0.817	1.505	-1.061	-0.554	na	-3.177	-3.038	na
dbscSNV_ada_score	na	na	na	na	na	0.999	0.999	na
dbscSNV_rf_score	na	na	na	na	na	0.926	0.934	na

Table S4. Cloning, mutagenesis and RT-PCR primers

DPF2_pEGFP_EcoRI_F	5´-CTCAAGCTTCGAATTGCGGCTGTGGTGGAGAAT -3´
DPF2_pEGFP_EcoRI_R	5´-GTCGACTGCAGAATTTCAAGAGGAGTTCTGGTTCTGGTA-3´
DPF2_pGEX4T1_BglII_F	5´-GGGCCCAGATCTTGGCGGCTGTGGTGGAGAATGTAG-3´
DPF2_pGEX4T1_ pCMV-Tag2b_SalI_R	5´-GGGCCCGTCGACTCAAGAGGAGTTCTGGTTCTG-3´
DPF2_pCMV-Tag2b_EcoRI_F	5´-GGGCCCGAATTCGCGGCTGTGGTGGAGAATGTAG-3´
DPF2_pGEX4T1_ pCMV-Tag2b_SalI_R	5´-GGGCCCGTCGACTCAAGAGGAGTTCTGGTTCTG-3´
DPF2_C276F_F	5´-CAACTACTGTGACTTCTTCCTGGGGGACTCAAAG-3´
DPF2_C276F_R	5´-CTTTGAGTCCCCCAGGAAGAAGTCACAGTAGTTG-3´
DPF2_C330W_F	5´-CGAGTGCAAATGTTGGAATATCTGCGGCAC-3'
DPF2_C330W_R	5´-GTGCCGCAGATATTCCAACATTTGCACTCG-3'
DPF2_R350H_F	5´-GTGATGACTGCGATCATGGCTACCACATGTAC-3´
DPF2_R350H_R	5´-GTACATGTGGTAGCCATGATCGCAGTCATCAC-3´
DPF2_Exon9_RT_F	5´-CAGTGCATCGAGTGCAAATGTTG-3´
DPF2_Exon11_RT_R	5´-CAGCTGCCAAAGGCAGGCTGGC-3´

Supplementary References

1. London, B., Michalec, M., Mehdi, H., Zhu, X., Kerchner, L., Sanyal, S., Viswanathan, P.C., Pfahnl, A.E., Shang, L.L., Madhusudanan, M., et al. (2007). Mutation in glycerol-3-phosphate dehydrogenase 1 like gene (GPD1-L) decreases cardiac Na+ current and causes inherited arrhythmias. Circulation 116, 2260-2268.
2. Bainbridge, M.N., Hu, H., Muzny, D.M., Musante, L., Lupski, J.R., Graham, B.H., Chen, W., Gripp, K.W., Jenny, K., Wienker, T.F., et al. (2013). De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. Genome medicine 5, 11.
3. Kuechler, A., Czeschik, J.C., Graf, E., Grasshoff, U., Huffmeier, U., Busa, T., Beck-Woedl, S., Faivre, L., Riviere, J.B., Bader, I., et al. (2017). Bainbridge-Ropers syndrome caused by loss-of-function variants in ASXL3: a recognizable condition. European journal of human genetics : EJHG 25, 183-191.
4. Kaasinen, E., Aittomaki, K., Eronen, M., Vahteristo, P., Karhu, A., Mecklin, J.P., Kajantie, E., Aaltonen, L.A., and Lehtonen, R. (2010). Recessively inherited right atrial isomerism caused by mutations in growth/differentiation factor 1 (GDF1). Human molecular genetics 19, 2747-2753.
5. Karkera, J.D., Lee, J.S., Roessler, E., Banerjee-Basu, S., Ouspenskaia, M.V., Mez, J., Goldmuntz, E., Bowers, P., Towbin, J., Belmont, J.W., et al. (2007). Loss-of-function mutations in growth differentiation factor-1 (GDF1) are associated with congenital heart defects in humans. American journal of human genetics 81, 987-994.
6. Watanabe, T.K., Fujiwara, T., Nakamura, Y., Hirai, Y., Maekawa, H., and Takahashi, E. (1996). Cloning, expression pattern and mapping to Xq of NAP1L3, a gene encoding a peptide homologous to human and yeast nucleosome assembly proteins. Cytogenetics and cell genetics 74, 281-285.
7. Labonne, J.D., Lee, K.H., Iwase, S., Kong, I.K., Diamond, M.P., Layman, L.C., Kim, C.H., and Kim, H.G. (2016). An atypical 12q24.31 microdeletion implicates six genes including a histone

demethylase KDM2B and a histone methyltransferase SETD1B in syndromic intellectual disability. Human genetics 135, 757-771.
8. Palumbo, O., Palumbo, P., Delvecchio, M., Palladino, T., Stallone, R., Crisetti, M., Zelante, L., and Carella, M. (2015). Microdeletion of 12q24.31: report of a girl with intellectual disability, stereotypies, seizures and facial dysmorphisms. American journal of medical genetics Part A 167A, 438-444.
9. Inoue, H., Furukawa, T., Giannakopoulos, S., Zhou, S., King, D.S., and Tanese, N. (2002). Largest subunits of the human SWI/SNF chromatin-remodeling complex promote transcriptional activation by steroid hormone receptors. The Journal of biological chemistry 277, 41674-41685.
10. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.
11. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research 22, 4673-4680.
12. Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics 31, 3997-3999.
13. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289-290.
14. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28-36.