The American Journal of Human Genetics, Volume *110*

Supplemental information

Monoallelic variation in *DHX9*, the gene encoding the DExH-box helicase DHX9, underlies neurodevelopment disorders and Charcot-Marie-Tooth disease

Daniel G. Calame, Tianyu Guo, Chen Wang, Lillian Garrett, Angad Jolly, Moez Dawood, Alina Kurolap, Noa Zunz Henig, Jawid M. Fatih, Isabella Herman, Haowei Du, Tadahiro Mitani, Lore Becker, Birgit Rathkolb, Raffaele Gerlini, Claudia Seisenberger, Susan Marschall, Jill V. Hunter, Amanda Gerard, Alexis Heidlebaugh, Thomas Challman, Rebecca C. Spillmann, Shalini N. Jhangiani, Zeynep Coban-Akdemir, Seema Lalani, Lingxiao Liu, Anya Revah-Politi, Alejandro Iglesias, Edwin Guzman, Evan Baugh, Nathalie Boddaert, Sophie Rondeau, Clothide Ormieres, Giulia Barcia, Queenie K.G. Tan, Isabelle Thiffault, Tomi Pastinen, Kazim Sheikh, Suur Biliciler, Davide Mei, Federico Melani, Vandana Shashi, Yuval Yaron, Mary Steele, Emma Wakeling, Elsebet Østergaard, Lusine Nazaryan-Petersen, Undiagnosed Diseases Network, Francisca Millan, Teresa Santiago-Sim, Julien Thevenon, Ange-Line Bruel, Christel Thauvin-Robinet, Denny Popp, Konrad Platzer, Pawel Gawlinski, Wojciech Wiszniewski, Dana Marafi, Davut Pehlivan, Jennifer E. Posey, Richard A. Gibbs, Valerie Gailus-Durner, Renzo Guerrini, Helmut Fuchs, Martin Hrabě de Angelis, Sabine M. Hölter, Hoi-Hung Cheung, Shen Gu, and James R. Lupski

Supplemental Notes

Limited clinical details were available for three individuals: two within the BHCMG/GREGoR, BAB4646 and M42-1, and an individual from a simplex autism spectrum disorder (ASD) cohort[19]. BAB4646 and M42-1 have the only two *DHX9* pLoF variants within the BHCMG database of 12,266 exomes and genomes. BAB4646's phenotype is severe DD/ID and primary immunodeficiency. Proband ES identified two heterozygous pathogenic variants in *TRNT1*(NM_182916.2): c.1246A>G, p.(Lys416Glu) and c.608+1G>T. As *TRNT1* causes autosomal recessive syndrome sideroblastic anemia with B-cell immunodeficiency, periodic fevers, and developmental delay [MIM: 616084], these variants likely contribute to the individual's DD/ID and immunodeficiency. Further confirmation of this contention was obfuscated by the lack of additional DNA samples from the proband or his parents precluding variant phasing and determination of *de novo* status.

M42-1 was enrolled in a mitochondrial disease cohort and is one of two affected siblings with encephalopathy, stroke-like episodes, and drug-resistant epilepsy. Proband ES failed to identify a candidate variant to explain the individual's mitochondrial disease, and sibling DNA is not available for testing. While *DHX9* pLoF variants are unlikely to completely explain the phenotypes of BAB4646 and M42-1, they may contribute to their neurologic dysfunction via multi-locus pathogenic variation to a blended traits phenotype[52].

		Age (weeks)									
Lab/Screen	Methods	7	8	9	10	11	12	13	14	15	16
Behaviour	Openfield		■								
	Acoustic startle response & PPI				■						
Neurology	Modified SHIRPA, grip strength			■							
	Rotarod				■						
Dysmorphology	Anatomical observation			■							
Energy Metabolism	Indirect calorimetry					■					
Cardiovascular	Awake ECG / Echo cardiography						■				
Clinical Chemistry	IpGTT							■			
Neurology	Auditory brain stem response (ABR)								■		
Dysmorphology	X-Ray, DEXA								■		
Eye	Scheimpflug imaging, Laser-interference-biometry (LIB), Optical coherence tomography (OCT), Virtual drum test									■	
Clinical Chemistry	Clinical Chemical analysis, hematology										■
Immunology	Flow cytometry, plasma (IgE, IL6, TNF, insulin)										■
Pathology	Macro & microscopic analysis										■

Figure S1 - Phenotyping pipeline used for *Dhx9*$^{-/-}$ mice
Mouse age in weeks for each phenotypic examination is highlighted in blue.

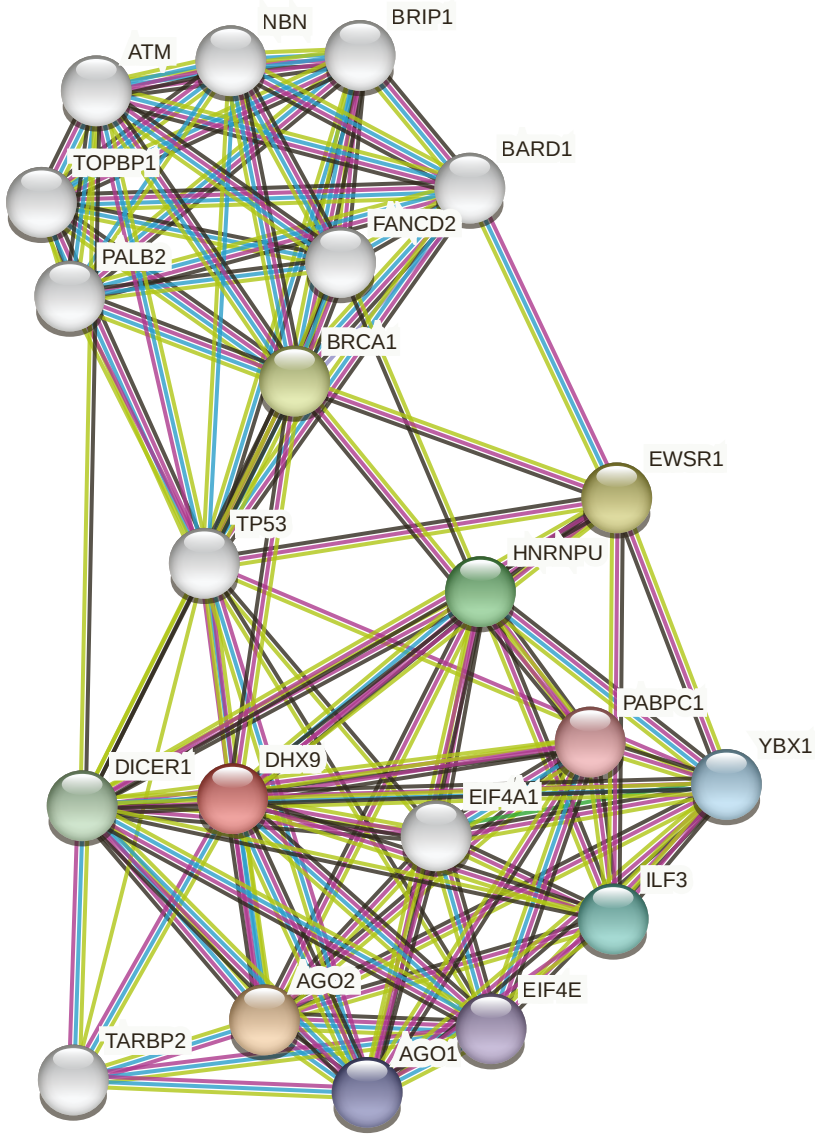


Figure S2 - DHX9 protein-protein interactions
DHX9 interactome data from STRING database (https://string-db.org/)

A)

	Nuclear localization signal																
H. sapiens	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
P. troglodytes	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
M. mulatta	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
P. hamadryas	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
B. taurus	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
E. caballus	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
C. lupus familiaris	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	G	G	Y
M. musculus	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
R. norvegicus	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
L. africana	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
D. novemcinctus	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	S	G	Y
A. carolinensis	Y	G	D	G	P	R	P	P	**K**	M	A	**R**	Y	D	N	G	G	G	Y
D. rerio	F	G	D	G	P	R	P	P	**K**	M	A	**R**	T	D	F	G	G	G	F

B) WT DHX9

Predicted NLSs in query sequence

```
IVLVDDWIKLQISHEAAACITGLRAAMEALVVEVTKQPAIISQLDPVNER 50
MLNMIRQISRPSAAGINLMIGSTRYGDGPRPPKMARYDNGSGYRRGGSSY 100
SGGGYGGGYSSGGYGSGGYGGSANSFRAGYGAGVGGGYRGVSRGGFRGNS 150
GGDYRGPSGGYRGSGGFQRGGGRGAYGTGYFGQGRGGGGY           190
```

Predicted monopartite NLS

Pos.	Sequence	Score
78	GPRPPKMARYD	7.5

C) p.(Lys1163Arg)

Predicted NLSs in query sequence

```
IVLVDDWIKLQISHEAAACITGLRAAMEALVVEVTKQPAIISQLDPVNER 50
MLNMIRQISRPSAAGINLMIGSTRYGDGPRPPRMARYDNGSGYRRGGSSY 100
SGGGYGGGYSSGGYGSGGYGGSANSFRAGYGAGVGGGYRGVSRGGFRGNS 150
GGDYRGPSGGYRGSGGFQRGGGRGAYGTGYFGQGRGGGGY           190
```

Predicted monopartite NLS

Pos.	Sequence	Score

D) p.(Arg1166Pro)

Predicted NLSs in query sequence

```
IVLVDDWIKLQISHEAAACITGLRAAMEALVVEVTKQPAIISQLDPVNER 50
MLNMIRQISRPSAAGINLMIGSTRYGDGPRPPKMAPYDNGSGYRRGGSSY 100
SGGGYGGGYSSGGYGSGGYGGSANSFRAGYGAGVGGGYRGVSRGGFRGNS 150
GGDYRGPSGGYRGSGGFQRGGGRGAYGTGYFGQGRGGGG            189
```

Predicted monopartite NLS

Pos.	Sequence	Score

Figure S3 - Analysis of nuclear localization missense variantsts

(A) Conservation of all amino acids which fall within the NLS. Lys1163 and Arg1166 are highlighted.
(B) cNLS Mapper prediction for reference sequence. Red letters indicate predicted NLS. Yellow highlight indicates known NLS.
(C) cNLS prediction for p.(Lys1163Arg) variant sequence. A NLS was not identified.
(D) cNLS prediction for p.(Arg1166Pro) variant sequence. A NLS was not identified.

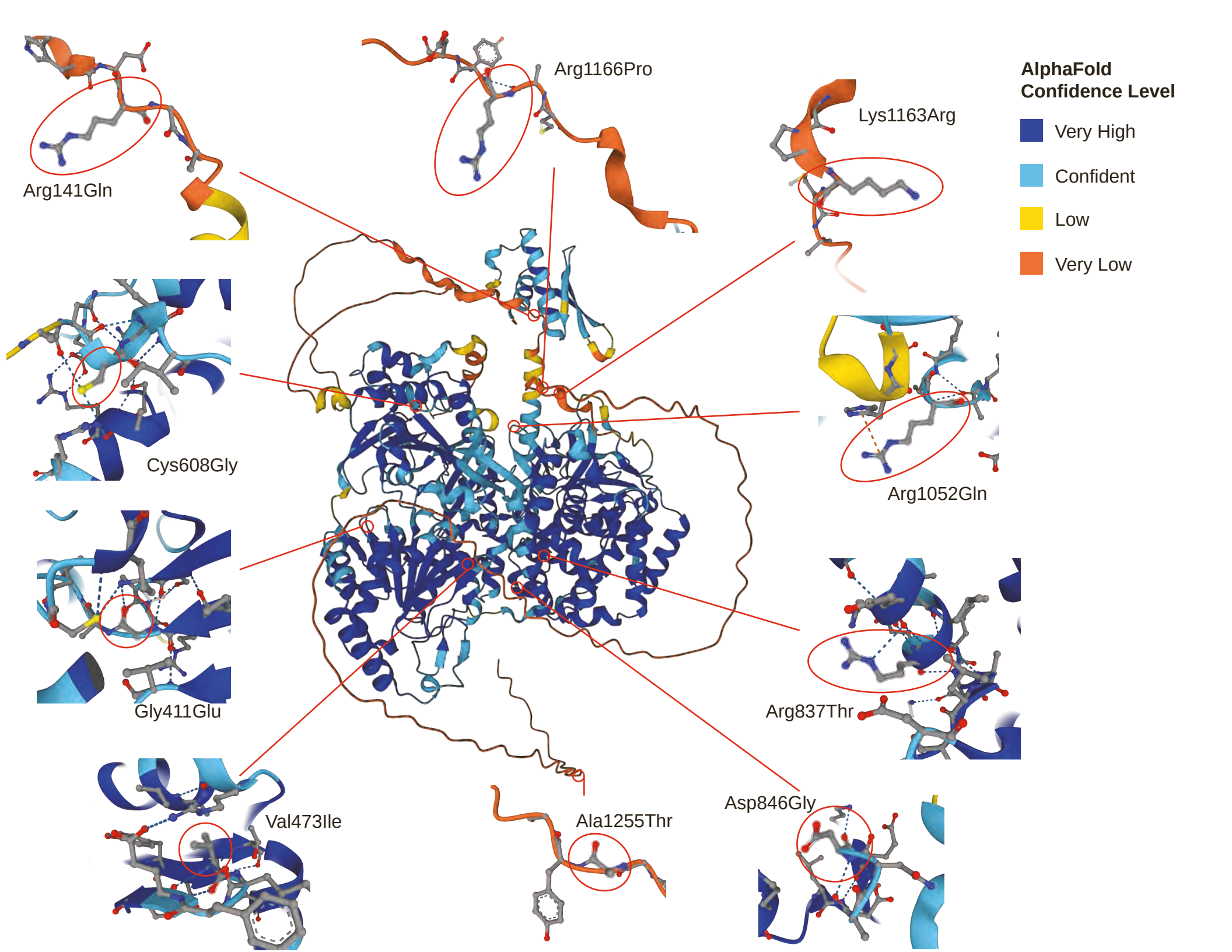
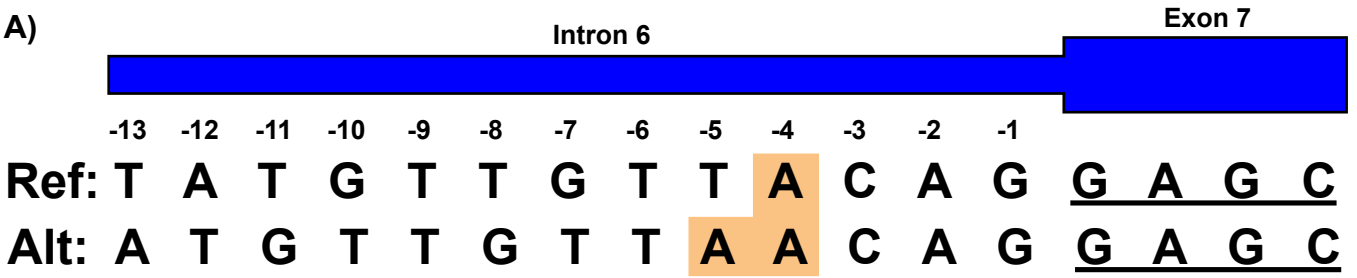

Arg141Gln
Arg1166Pro
Lys1163Arg
AlphaFold
Confidence Level
Very High
Confident
Low
Very Low
Cys608Gly
Arg1052Gln
Gly411Glu
Arg837Thr
Val473Ile
Ala1255Thr
Asp846Gly

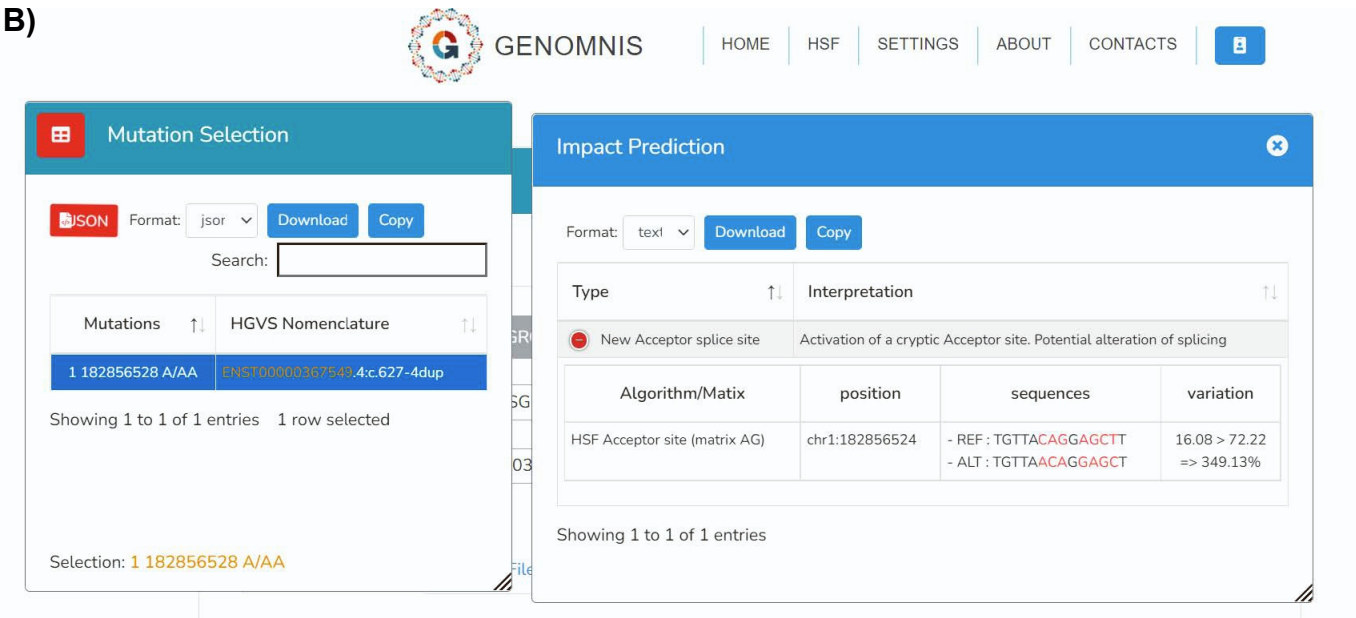
Figure S4 - Visualization of DHX9 variant alleles with 3D AlphaFold structure of DHX9

The AlphaFold predicted structure of DHX9 (UniProt: Q08211) is shown at center. The amino acids affected by reported missense variants are highlighted by a red circle in the whole protein structure modeled at center, and by a red circle or ellipse at the periphery in local structures at the periphery. The confidence level of AlphaFold predicted structure is denoted by color, and a key representing the per-residue confidence level for each color is shown at right[4,5].

A)



B)



C)

Δ type	Δ score ⓘ	pre-mRNA position ⓘ
Acceptor Loss	0.02	4 bp
Donor Loss	0.01	50 bp
Acceptor Gain	0.14	18 bp
Donor Gain	0.00	

Figure S5 - *In silico* analysis of *DHX9* variant c.627-4dupA

(A) Diagram demonstrating impact of *DHX9*: c.627-4dupA on splicing junction.
(B) *In silico* prediction from Human Splice Finder (http://www.umd.be/hsf/)
(C) *In silico* prediction from Splice AI (https://spliceailookup.broadinstitute.org/).up.broadinstitute.org/).

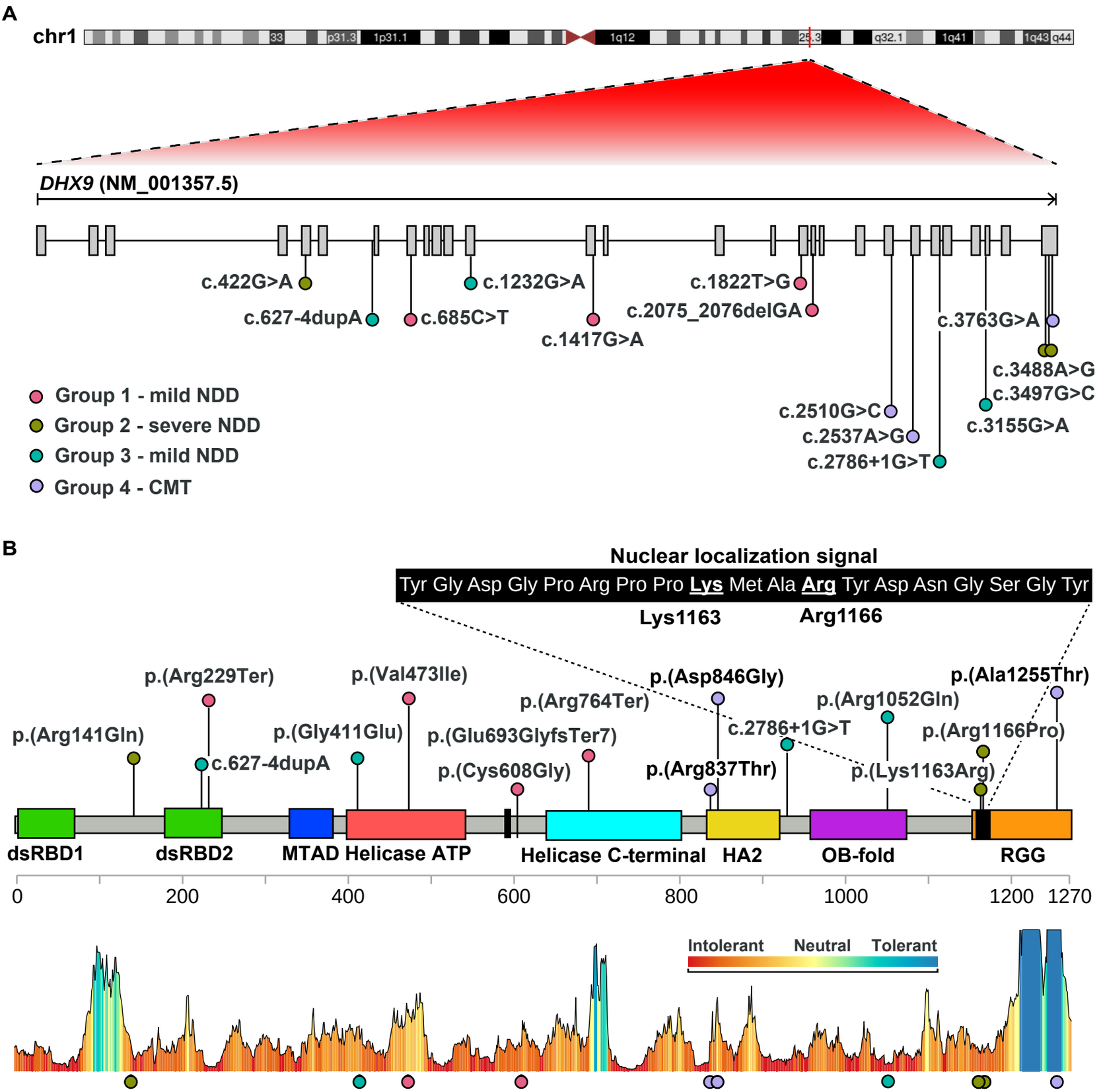


Figure S6 - Visualization of *DHX9* variant alleles used in HPO analysis

(A) Diagram of DHX9 mRNA showing location of DHX9 variants used in HPO analysis. Group 1 (mild NDD) = pink, Group 2 (severe NDD) = green, Group 3 (mild NDD) = teal, Group 4 (CMT) = purple.

(B) Diagram of DHX9 protein showing functional domains including double-stranded RNA-binding domains (dsRBD1&2), minimal transactivation domain (MTAD), helicase domains, helicase associated domain 2 (HA2), oligonucleotide/oligosaccharide-binding fold (OB-fold), and the RGG box. Protein domains were obtained from Uniprot. The sequence of the nuclear localization signal is

magnified and the two key residues Lys1163 and Arg1166 are underlined. DHX9’s protein tolerance landscape is shown below the figure as calculated by Metadome.

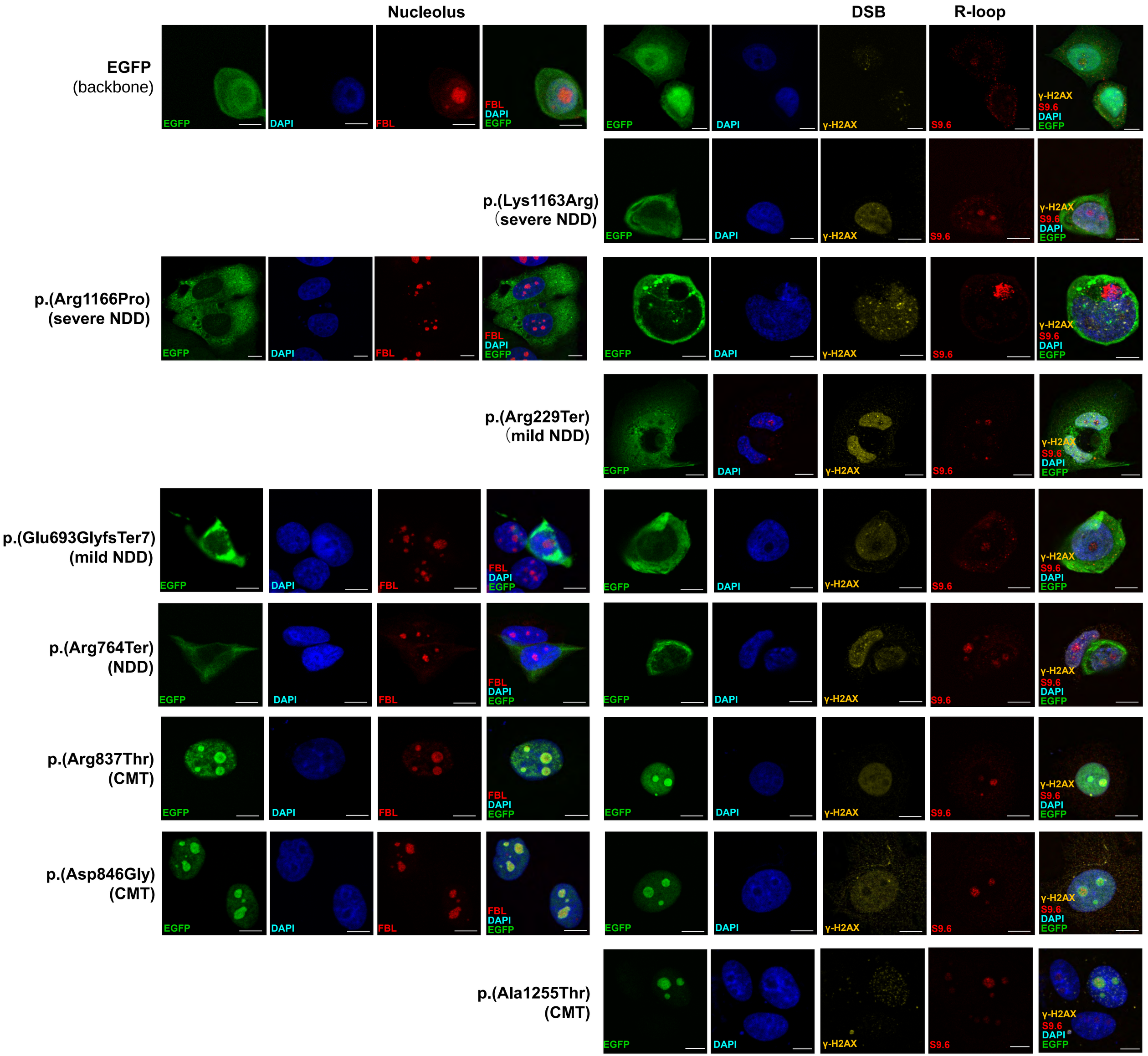


Figure S7 - Subcellular localization of DHX9 variant proteins and the levels of R-loop and DNA damage
Scale bar = 10 μm. Subcellular localization of EGFP-tagged DHX9 variant proteins. Nucleolar loci were co-stained by the fibrillarin (FBL) marker and DNA stained by DAPI. Staining of levels of R-loop formation by the S9.6 marker and DSB by the γ-H2AX marker.

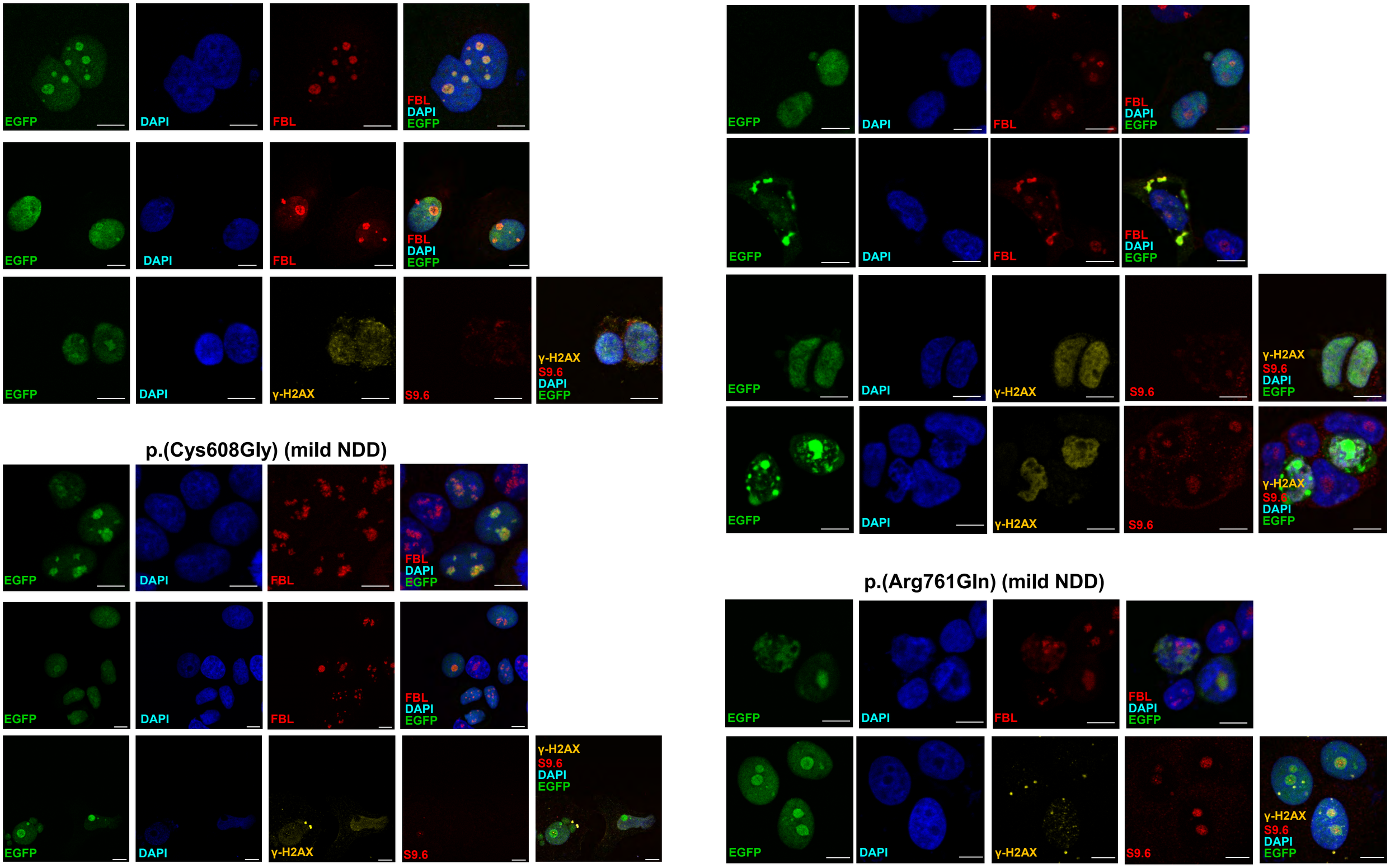


Figure S8 - Subcellular localization of remaining DHX9 variant proteins and the levels of R-loop and DNA damage
Scale bar = 10 μm. Subcellular localization of EGFP-tagged DHX9 variant proteins. Nucleolar loci were co-stained by the fibrillarin (FBL) marker and DNA stained by DAPI. Staining of levels of R-loop formation by the S9.6 marker and DSB by the γ-H2AX marker.

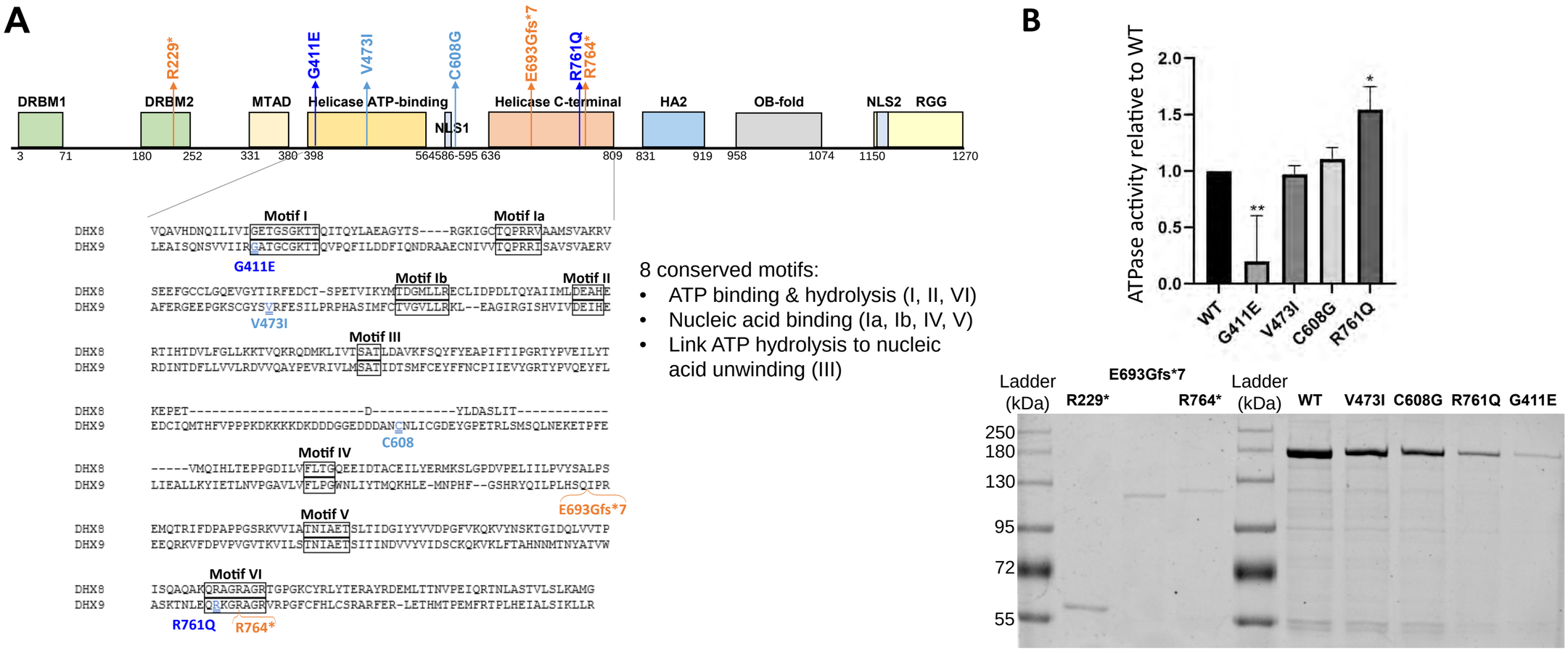
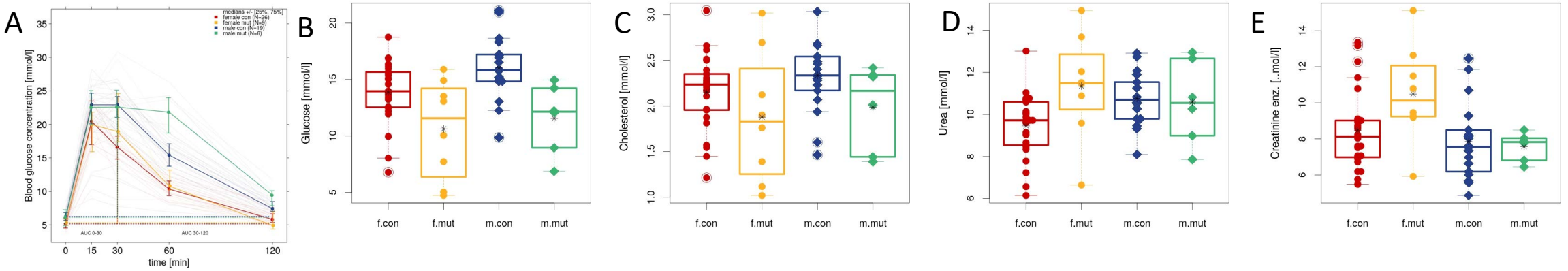


Figure S9 - *DHX9* missense variants located within the ATP binding and hydrolysis conserved motifs affected ATPase activity.

(A) Schematic representation of patients' variants with regards to functional domains of the DHX9 protein. Amino acid sequence of the helicase ATP-binding and helicase C-terminal domains of DHX9 is listed. The eight conserved motifs of these functional domains are marked within boxes. Corresponding sequence of DHX8 is aligned together to demonstrate the high conservation of the eight motifs. Truncating variants (R229*, E693Gfs*7 and R764*) are labeled in orange, missense variants located within conserved motifs (G411E and R761Q) are labeled in dark blue, and the remaining missense variants (V473I and C608G) are labeled in light blue.

(B) Assays demonstrated the relative ATPase activities of DHX9 variant proteins with various missense changes compared to WT protein. In each experiment, ATPase activity was normalized to the amount of purified protein. A representative Coomassie blue staining image demonstrating the sizes and expression levels of purified DHX9 proteins is shown here. Experiments were repeated at least three times for each DHX9 variant. See Table S2 for raw data on absorbance values of each sample. Note that for truncating variants (R229*, E693Gfs*7 and R764*), the absorbance values were comparable to the baseline (no transfection blank and EGFP backbone only expression), therefore, their relative ATPase activities to WT protein were not calculated. Also note that for the p.R761Q protein, its much lower expression level relative to WT caused its higher calculated ATPase activity, given the calculation was normalized based on corresponding protein amount. Its actual ATPase activity values (raw data on absorbance values) were consistently lower than the WT values (Table S2). **, $p<0.005$; *, $p<0.05$; One-Way ANOVA.



IpGTT results	female		male		linear model	linear model	linear model
	con	mut	con	mut	genotype	sex	genotype:sex
	n=26	n=9	n=19	n=6			
	mean ± sd	mean ± sd	mean ± sd	mean ± sd	p-value	p-value	p-value
Glucose (T=0)	5.35 ± 0.81	5.42 ± 0.56	6.23 ± 1.04	6.03 ± 1.27	0.822	0.014	0.664
AUC 0-30	295.39 ± 94.35	330.05 ± 74.5	366.05 ± 51.65	384.82 ± 65.12	0.302	0.018	0.758
AUC 30-120	417.75 ± 176.08	551.57 ± 208.3	724.27 ± 183.29	1050.96 ± 226.25	< 0.001	< 0.001	0.123

Figure S10 - Loss of *Dhx9* in mice causes differences in clinical chemistry indices indicative of altered metabolism and renal function.

Results of blood chemistry tests are compared between control (WT) mice and mutant mice (*Dhx9*$^{-/-}$). Red = female controls, yellow = female mutants, blue = male controls, green = male mutants.

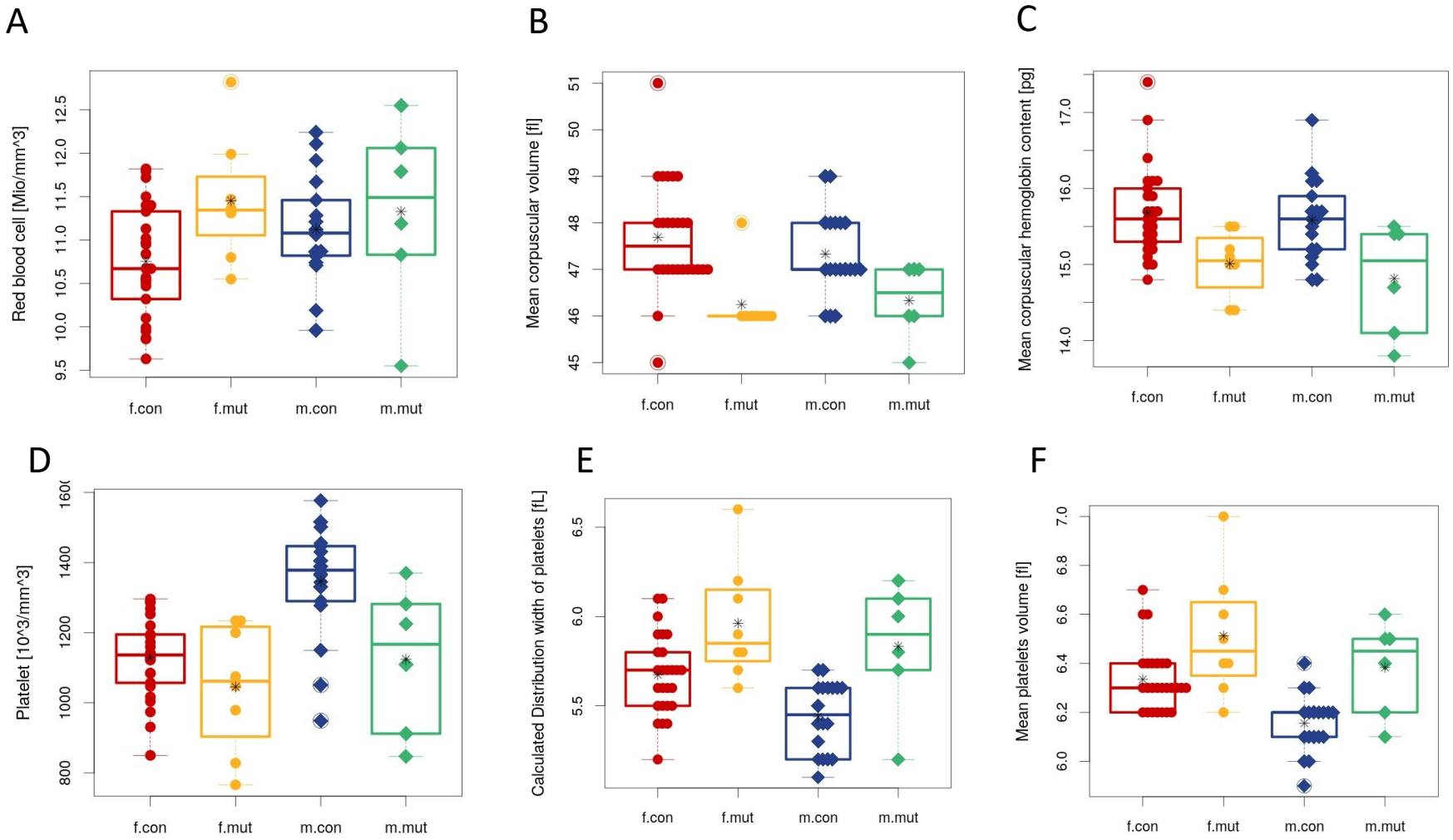


Figure S11 - Loss of *Dhx9* in mice causes hematological alterations indicative of effects on erythropoiesis and thrombopoiesis

Hematological testing comparing control (WT) mice and mutant mice (*Dhx9*$^{-/-}$). * indicates statistical significance ($p<0.05$). Red = female controls, yellow = female mutants, blue = male controls, green = male mutants.

Assay	Age (weeks)	Number (n)			
		+/+		-/-	
		Males	Females	Males	Females
Open field	8	19	26	6	9
SHIRPA	9	19	26	6	9
Grip strength	9	19	26	6	9
Acoustic startle	10	19	26	6	9
Indirect calorimetry	11	18	26	6	9
Glucose tolerance test	13	19	26	6	9
Auditory brainstem response	14	14	16	4	4
Clinical chemistry/hematology	16	19	26	6	9

Table S1 - Number of *Dhx9*$^{-/-}$ mice tested in the assays where relevant differences were detected..

Individual	1	2	3	4	5	6	7	8	9	10	11	12
Sex	M	F	M	M	F	M	M	F	F	M	M	M
Age at last examination	16 y	16 y	5 y	3.5 y	19 y	8 y	7 y	15 y	11 y	23 y	3 y	12 y
Phenotype	NDD	NDD	NDD	NDD	NDD	NDD	NDD	NDD	NDD	NDD	NDD	NDD
Developmental delay	+	+	+	+	+	+	+	+	+	+	+	+
Intellectual disability	Severe	Borderline	Severe	n/a	-	Severe	Mild	Mild	Mild	-	n/a	Severe
Microcephaly (Z-score)	+ (-2.14)	-	+ (-2.49)	-	-	+ (-3.16)	-	-	-	-	-	+ (-3.22)
Abnormal brain MRI	+	-	+	-	-	+	-	-	-	n.d.	-	+
Neuropsychiatric disorders	-	+	+	+	+	-	+	+	+	-	+	-
Seizures	+	-	-	-	+	-	-	+	+	-	-	-
Drug-resistant epilepsy	+	-	-	-	+	-	-	-	+	-	-	-
Axial hypotonia	+	+	+	-	+	+	-	-	-	-	-	+
Appendicular hypertonia	+	-	-	-	-	+	-	-	-	-	-	-
Abnormal reflexes	Incr.	-	Incr.	-	-	Incr.	-	-	-	-	-	-
Ataxia	-	-	+	-	-	-	-	-	-	-	-	+
Axonal neuropathy	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	-	n.r.	n.r.
Dysmorphic features	-	+	+	-	-	+	+	+	+	-	-	+
Heart disease	-	+	+	-	+	-	-	-	-	+	-	-
Short stature	+	-	-	-	-	+	-	-	-	-	-	-
Failure to thrive	+	-	-	-	-	+	-	-	-	-	-	-
Recurrent infections	+	-	-	-	-	+	-	-	-	-	-	-

Abbreviations: M, male; F, female; NDD, neurodevelopmental disorder; CMT, Charcot-Marie-Tooth disease; n.r., not reported; n.d., not done.; n/a, not applicable; Incr., increased; Dim., diminished.; Mod, moderate.

Individual	13	14	15	16	17	Total NDD	Total All
Sex	F	M	F	M	M	-	-
Age at last examination	8 y	6 mo	45 y	54 y	58 y	-	-
Phenotype	NDD	NDD	CMT	CMT	CMT	-	-
Developmental delay	+	+	-	-	-	14/14	14/17
Intellectual disability	Mod.	n/a	-	-	-	8/11	8/14
Microcephaly (Z-score)	+ (-2.3)	+ (-3.39)	-	-	-	6/14	6/17
Abnormal brain MRI	+	n.d.	n.d.	n.d.	n.d.	5/12	5/12
Neuropsychiatric disorders	-	-	-	-	-	8/14	8/17
Seizures	+	+	-	-	-	6/14	6/17
Drug-resistant epilepsy	-	-	-	-	-	3/14	3/17
Axial hypotonia	-	+	-	-	-	7/14	7/17
Appendicular hypertonia	-	-	-	-	-	2/14	2/17
Abnormal reflexes	-	-	Dim.	Dim.	Dim.	3/14	6/17
Ataxia	-	-	+	-	-	2/14	3/17
Axonal neuropathy	n.r.	n.r.	+	+	+	1/14	4/17
Dysmorphic features	+	-	-	-	-	8/14	8/17
Heart disease	-	-	-	-	-	4/14	4/17
Short stature	+	-	-	-	-	3/14	3/17
Failure to thrive	-	-	-	-	-	2/14	2/17
Recurrent infections	-	-	-	-	-	2/14	2/17

Abbreviations: M, male; F, female; NDD, neurodevelopmental disorder; CMT, Charcot-Marie-Tooth disease; n.r., not reported; n.d., not done.; n/a, not applicable; Incr., increased; Dim., diminished.; Mod, moderate.

Table S2 - Phenotypic Summary of Individuals with *DHX9*-related Neurodevelopmental Disorders and CMT

	Absorbance at 620nm			
Transfected expression plasmid	Trial 1	Trial 2	Trial 3	Average
No transfection blank	0.442	0.441	0.441	0.441
EGFP-backbone	0.435	0.342	0.343	0.373
EGFP-DHX9 WT	0.646	0.647	0.66	0.651
DHX9 p.(Arg229Ter)	0.377	0.433	0.417	0.409
DHX9 p.(Gly411Glu)	0.457	0.419	0.466	0.447
DHX9 p.(Val473Ile)	0.628	0.653	0.634	0.638
DHX9 p.(Cys608Gly)	0.648	0.688	0.656	0.664
DHX9 p.(Glu693GlyfsTer7)	0.438	0.358	0.421	0.406
DHX9 p.(Arg764Ter)	0.41	0.447	0.428	0.428
DHX9 p.(Arg761Gln)	0.6	0.614	0.593	0.602

Table S3 - Raw data on absorbance values of ATPase activity experiments.

Supplemental References

1. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of *de novo* coding mutations to autism spectrum disorder. Nature *515*, 216–221. 10.1038/nature13908.

2. Mitani, T., Isikay, S., Gezdirici, A., Gulec, E.Y., Punetha, J., Fatih, J.M., Herman, I., Akay, G., Du, H., Calame, D.G., et al. (2021). High prevalence of multilocus pathogenic variation in neurodevelopmental disorders in the Turkish population. Am. J. Hum. Genet. *108*, 1981–2005. 10.1016/j.ajhg.2021.08.009.

3. Herman, I., Jolly, A., Du, H., Dawood, M., Abdel-Salam, G.M.H., Marafi, D., Mitani, T., Calame, D.G., Coban-Akdemir, Z., Fatih, J.M., et al. (2022). Quantitative dissection of multilocus pathogenic variation in an Egyptian infant with severe neurodevelopmental disorder resulting from multiple molecular diagnoses. Am. J. Med. Genet. A. *188*, 735–750. 10.1002/ajmg.a.62565.

4. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. 10.1038/s41586-021-03819-2.

5. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. *50*, D439–D444. 10.1093/nar/gkab1061.

Supplemental Acknowledgements

This study was supported by the CHU de Dijon Bourgogne and by a grant from the French Ministry of Health «DIS-SEQ - Evaluation médico-économique des différentes stratégies de technologies de séquençage par haut débit dans le diagnostic des patients atteints de déficience intellectuelle», Clinical Trial NCT03287206. The Meyer Children's Hospital IRCCS is a full Member of the ERN Epicare. This study was also supported by Regione Toscana under the Call for Health 2018 (grant DECODE-EE) (to Re.Gu.) and the 'Brain Project' by Fondazione Cassa di Risparmio di Firenze (to Re.Gu.). M.H.d.A. was funded by German Federal Ministry of Education and Research (Infrafrontier grant 01KX1012) and the German Center for Diabetes Research (DZD). The Deciphering Developmental Disabilities (DDD) study presents

independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. The work by Tomi Pastinen and Isabelle Thiffault was made possible by the generous gifts to Children's Mercy Research Institute and Genomic Answers for Kids program at Children's Mercy Kansas City. Tomi Pastinen holds the Dee Lyons/Missouri Endowed Chair in Pediatric Genomic Medicine. Research reported in this manuscript was supported by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under Award Number(s) [UO1HG007672- Shashi].