The power of the Mediator complex - expanding the genetic architecture and phenotypic spectrum of *MED12*-related disorders

Running title: **A spectrum of *MED12*-related disorders**

Agnieszka Charzewska[1*], Robert Maiwald[2], Kimia Kahrizi[3], Barbara Oehl-Jaschkowitz[4], Andreas Dufke[5], Johannes R. Lemke[5,6], Herbert Enders[5,7], Hossein Najmabadi[3], Andreas Tzschach[8], Wiebke Hachmann[9], Corinna Menzel[10], Melanie Bienek[10], Jarosław Poznański[12], Magdalena Nawara[1], Tatiana Chilarska[11], Ewa Obersztyn[1], Dorota Hoffman-Zacharska[1], Monika Gos[1], Jerzy Bal[1], and Vera M Kalscheuer[13*]

[1]Institute of Mother and Child, Department of Medical Genetics, Warsaw, Poland

[2]MVZ für Medizinische Genetik und Molekulare Medizin, Cologne, Germany

[3] Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[4]Gemeinschaftspraxis für Humangenetik, Biomedizinisches Zentrum, Homburg Saar, Germany

[5]Institut für Medizinische Genetik und Angewandte Genomik, Tübingen, Germany

[6]Institut für Humangenetik am Universitätsklinikum Leipzig AöR, Germany

[7]MVZ Humangenetik Ulm, Germany

[8]Institut für Klinische Genetik, Medizinische Fakultät CGC, Technische Universität Dresden, Germany, and Institut für Medizinische Genetik und Angewandte Genomik, Tübingen, Germany

[9]Elisabeth-Krankenhaus Rheydt, Klinik für Kinder und Jugendliche, Mönchengladbach, Germany

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cge.13412

[10]Max Planck Institute for Molecular Genetics, Department of Human Molecular Genetics, Berlin, Germany

[11]Polish Mother's Memorial Hospital Research Institute, Department of Genetics, Łódź, Poland

[12]Polish Academy of Sciences, Institute of Biochemistry and Biophysics, Department of Biophysics, Warsaw, Poland

[13]Max Planck Institute for Molecular Genetics, Research Group Development and Disease, Berlin, Germany

*Corresponding authors:

Agnieszka Charzewska, Ph.D.

Institute of Mother and Child

Department of Medical Genetics

Kasprzaka 17 A, 01-211 Warsaw, Poland

E-mail: agnieszka.charzewska@imid.med.pl

Phone: +48 22 32 77 313

Fax: +48 22 32 77 200

Vera M. Kalscheuer, Ph.D.

Max Planck Institute for Molecular Genetics

Research Group Development and Disease

Ihnestr. 73, D-14195 Berlin, Germany

Phone: (49)30 8413 1293

Fax: (49) 30 8413 1383

E-mail: kalscheu@molgen.mpg.de

Webpage: www.molgen.mpg.de/3415877/Chromosome-Rearrangements-and-Disease

ACKNOWLEDGEMENTS

This work was supported by Polish National Science Centre grant number 2014/15/D/NZ5/03426 (ACh) and by the EU FP7 project GENCODYS, grant number 241995 (VMK).

The authors declare that they have no conflict of interest.

ABSTRACT

MED12 is a member of the large Mediator complex that controls cell growth, development, and differentiation. Mutations in *MED12* disrupt neuronal gene expression and lead to at least three distinct X-linked intellectual disability (XLID) syndromes (FG, Lujan-Fryns, and Ohdo). Here we describe six families with missense variants in *MED12* (p.(Arg815Gln), p.(Val954Gly), p.(Glu1091Lys), p.(Arg1295Cys), p.(Pro1371Ser) and p.(Arg1148His), the latter being firstly reported in affected females) associated with a continuum of symptoms rather than distinct syndromes. The variants expanded the genetic architecture and phenotypic spectrum of *MED12*-related disorders. New clinical symptoms included brachycephaly, anteverted nares, bulbous nasal tip, prognathism, deep set eyes, and single palmar crease. We showed that *MED12* variants, initially implicated in X-linked recessive disorders in males, may predict a potential risk for phenotypic expression in females, with no correlation of the X chromosome inactivation pattern in blood cells. Molecular modeling (Yasara Structure) performed to model the functional effects of the variants strongly supported the pathogenic character of the variants examined. We demonstrated that molecular modeling is a useful method for *in silico* testing of potential functional effects of *MED12* variants and thus can be a valuable addition to the interpretation of the clinical and genetic findings.

KEY WORDS

MED12, X-linked intellectual disability, FG syndrome, Lujan-Fryns syndrome, Ohdo syndrome, molecular modeling

INTRODUCTION

The Mediator is a large multiprotein complex that regulates gene expression in all eukaryotes and interacts with RNA polymerase II. The role of the complex involves transcriptional elongation and termination, mRNA processing, noncoding RNA activation, super enhancer formation, and epigenetic regulation. Thus, the Mediator was called an ‘‘integrative hub’’ for transcriptional processes and emerged as a master coordinator of cell growth and homeostasis, development, and differentiation[1]. The Mediator complex (MED) consists of 31 subunits (MED1-MED31) and can be divided into four distinct modules termed as the head, middle, tail, and CDK8 kinase module. The latter module contains CDK8, cyclin C, MED12 and MED13. MED12 is a critical transducer of regulatory information essential for organogenesis. At least three different intellectual disability (ID) conditions that are associated with *MED12* mutations have been described. These conditions include Opitz-Kaveggia syndrome (FG syndrome type 1, FGS1), Lujan-Fryns syndrome (LFS), and X-linked Ohdo syndrome (XLOS, OHDOX)[2]. These syndromes are allelic disorders that share clinical findings including ID, hypotonia and some physical features, such as tall prominent forehead, open mouth or high narrow palate. Probands with non-syndromic X-linked ID (XLID), including females with variable cognitive impairment, have also been described as harboring *MED12* mutations. Here, we present clinical and genetic details on six families with missense variants in the *MED12* gene, as well as molecular modeling results for four variants, and expand the phenotypic spectrum of *MED12*-related disorders.

MATERIALS AND METHODS

Six ID families with *MED12* variants (NM_005120) were included in the study (Fig.1). Patients' phenotypes are described in Supplementary material 1 and summarized in Supplementary material 2. Families 1 to 4 were collected as part of the X-exome resequencing project of unresolved families with assumed XLID, while Families 5 and 6 were sequenced in routine diagnostics. For none of the families a clinical suggestion had been made prior to next generation sequencing (NGS). The procedures employed were reviewed and approved by the appropriate institutional review committees. DNAs were isolated from peripheral blood according to standard procedures. Written informed consent was obtained for all individuals participating in this study. For the index probands NGS was performed in three laboratories (Berlin, Tübingen, and Warsaw) as described in Supplementary material 3. Segregation analysis was performed by Sanger sequencing for the other available family members. The pathogenicity of variants obtained from NGS was assessed by molecular modeling (Yasara Structure) as described in Supplementary material 3.

RESULTS

We report on 13 affected males and 2 affected females from 6 families with variants in the *MED12* gene (Fig. 1). Clinical symptoms of the affected males and females along with a comparison with symptoms described in published *MED12*-related syndromes are presented in Table 1. All probands had ID and developmental delay, most of them had macrocephaly, long narrow face, prominent forehead, small ears, dental abnormalities, prognathism and hypertelorism. The remaining clinical features were variable. The new clinical findings associated with *MED12* variants (present in at least 2 families) were as follows: brachycephaly, anteverted nares, bulbous nasal tip, prognathism, deep set eyes, and single palmar crease. In the families we identified 3 novel missense variants in *MED12* (p.(Val954Gly), p.(Glu1091Lys), p.(Pro1371Ser)), 1 known variant but firstly reported in affected females (p.(Arg1148His)), as well as 2 variants (p.(Arg815Gln) and p.(Arg1295Cys)) which have already been described as a cohort data by the authors of the current paper, but without clinical details of the presented families[3,4].

All *MED12* variants identified in the families perfectly co-segregated with the phenotype, except for Family 5 in whom the *MED12* variant present in the affected daughters was inherited from the healthy mother. Most of the variants were absent in the gnomAD database (except for c.2444G>A present in 1 heterozygous female with a minor allele frequency (MAF) of 5.599e-6 and c.4111C>T present in 6 hemizygous individuals with a MAF of 0.00012). All the variants were predicted as damaging by Polyphen2, MutationTaster, as well as SIFT, except for c.4111C>T identified in patient III:1 from Family 6 which had conflicting interpretations of pathogenicity (Polyphen2 – benign, Mutation Taster – disease causing, SIFT – tolerated). Additionally, this variant had been reported to ClinVar as likely benign. NGS performed subsequently in the patient revealed additionally a *de novo* heterozygous missense variant c.367C>T in the *PUF60* gene

(NM_078480.2) resulting in the p.(Arg123Trp) substitution. The variant was predicted as damaging (Polyphen2, MutationTaster, SIFT) and it was absent in gnomAD database.

Molecular modeling was applied to assess the functional effects of the *MED12* (four of the six identified MED12 substitutions are located within the protein regions, the structure of which may be reasonably modeled by homology) and *PUF60* variants identified by NGS. All the variants are expected to have substantial structural effects on proteins (Fig. 2).

X chromosome inactivation (XCI) analysis performed in Family 4 revealed that the mother (II:1) of the index patient (III:2) who also carried the mutation had mildly skewed XCI (83:17) in her blood cells. In Family 5 XCI analysis revealed a mildly skewed pattern in the mother (82:18) and a skewed pattern in her two affected daughters (100:0 and 85:15). In the three females, the same allele of the AR locus was inactive.

DISCUSSION

The phenotypic spectrum of *MED12*-related disorders is still being expanded as new families with *MED12* variants are being identified due to the improvement of new sequencing technologies. Moreover, it was postulated that succeeding syndromes should be defined to extend the list of the three allelic syndromes associated with *MED12* mutations. However, as shown by the clinical data of our probands (Table 1) it seems that they do not manifest strict syndromes (FGS1, Lujan-Fryns or Ohdo), but they rather present, similar to recently published data[5], a continuum of symptoms. Thus, it is more appropriate to indicate only a '*MED12*-related disorder' than to attribute a definite syndrome for a majority of the subjects.

To date, over 20 *MED12* genetic variants have been reported which are evenly distributed with no apparent hot spots in specific exons (Fig. 1). All but one were missense variants causing FGS1, Lujan-Fryns, Ohdo syndrome or non-syndromic ID with different degrees of cognitive impairment. Many benign variants (456 in ClinVar) were also described as well as somatic variants implicated in cancer but not associated with germline mutations. The *MED12* variants examined in the current study were defined as pathogenic by molecular modeling, the approach that is not only based on bioinformatic prediction, but also uses experimentally gained protein crystal structures. The results suggested a pathogenic effect of all four *MED12* variants tested: p.(Val954Gly), p.(Glu1091Lys), p.(Arg1295Cys), and p.(Pro1371Ser). The other two variants (p.(Arg815Gln) and p.(Arg1148His)) were not covered by the models obtained. Despite these results, the maternally inherited *MED12* p.(Pro1371Ser) substitution identified in the sporadic male (Family 6) and also present in hemizygous individuals in gnomAD was found to be problematic in the case of the *MED12* variant's pathogenicity. Because the *MED12* variant was found in the proband by X-chromosome exome re-sequencing, we performed subsequent WES in order to exclude the possibility of other variants being engaged in ID pathogenesis. The WES analysis revealed a

de novo p.(Arg123Trp) variant in the *PUF60* gene. *PUF60* variants cause autosomal dominant Verheij syndrome which is clinically highly variable. Main clinical symptoms include ID, short stature, cerebral atrophy, microcephaly, as well as renal, skeletal, and cardiac defects[6]. With the current knowledge, the *de novo PUF60* variant identified in this study only partially explains the phenotype of the proband from Family 6. The additional clinical findings such as macrocephaly, ptosis, Hirschsprung disease, genital hypoplasia, and anal anomaly fit in the clinical spectrum of *MED12*-related disorders and to the best of our knowledge have not been reported in probands with *PUF60* variants so far. Both, the *MED12* and *PUF60* variants were predicted to be pathogenic by molecular modeling and combined they better explain the phenotype than each individually. Thus, it is possible that the proband from Family 6 presents a complex phenotype caused by these two variants. Such digenic inheritance has recently been postulated as an important cause of ID.

Currently, reports on *MED12*-related clinical phenotypes in females are rather scarce. With the current knowledge, pathogenicity of the firstly reported *MED12* p.(Arg1148His) substitution in affected females (Family 5) and assumed as causing Ohdo syndrome in males only, by clinical, genetic and functional findings remains to be established[7]. Segregation analysis revealed that the variant was inherited from their healthy mother. Although *MED12* is subject to X-inactivation, there was no clear correlation between clinical phenotypes of the *MED12* carrier females from Family 5 and their XCI pattern in blood cells, similar to the results published previously[8]. It seems that blood DNA-based X inactivation does not predict clinical outcome in *MED12*-related carrier females and different affection statuses in females might be related to varying X-inactivation in the brain. Thus, it is possible that similar to other genes initially implicated in X-linked recessive disorders in males, missense variants in *MED12* may predict a potential risk for phenotypic expression in carrier females. This

assumption needs to be more firmly established in the future, and if confirmed would be important for genetic counseling and prognostic evaluation.

REFERENCES

[1] Yin J, Wang G. The Mediator complex: a master coordinator of transcription and cell lineage development. Development 2014;141:977-987.

[2] Lyons MJ. MED12-Related Disorders. In: Pagon RA, Adam MP, Ardinger HH, et al. GeneReviews, University of Washington, Seattle; 2008;1993-2016.

[3] Tzschach A, Grasshoff U, Beck-Woedl S, et al. Next-generation sequencing in X-linked intellectual disability. Eur J Hum Genet 2015;23(11):1513-1518.

[4] Hu H, Haas SA, Chelly J, et al. X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. Mol Psychiatry 2016;21(1):133-48.

[5] Langley KG, Brown J, Gerber RJ, et al. Beyond Ohdo syndrome: A familial missense mutation broadens the MED12 spectrum. Am J Med Genet A 2015;167A(12):3180-3185.

[6] Verheij JB, de Munnik SA, Dijkhuizen T, et al. An 8.35 Mb overlapping interstitial deletion of 8q24 in two patients with coloboma, congenital heart defect, limb abnormalities, psychomotor retardation and convulsions. Europ J Med Genet 2009;52:353-357.

[7] Vulto-van Silfhout AT, de Vries BB, et al. Mutations in MED12 cause X-linked Ohdo syndrome. Am J Hum Genet 2013;92(3):401-406.

[8] Lesca G, Moizard MP, Bussy G, et al. Clinical and neurocognitive characterization of a family with a novel MED12 gene frameshift mutation. Am J Med Genet A 2013;161A(12):3063-3071.

FIGURE LEGENDS

Fig. 1

A Pedigrees of the families with segregation of the *MED12* variants (black) and the *PUF60* variant (blue). Mutant alleles are represented by a plus (+) and wild-type alleles by a minus (–). **B** The distribution of the *MED12* variants in relation to the gene structure consisting of 45 exons and three functional domains: MED12, LCEWAV, and catenin-binding domain – PQL (domain positioning according to Pfam 31.0 database). Known pathogenic *MED12* variants are indicated above the gene and novel variants below the gene. Asterisks denote *MED12* variants present in affected females. **C** Images of *MED12* probands from the presented families. Most of the probands presented with a long narrow face, prominent forehead, prominent nasal bridge, small ears, and hypertelorism. Novel facial dysmorphisms that expand the phenotypic spectrum of *MED12*-related disorders include deep set eyes (Families 1, 5), prognathism (Families 1, 3, 5, 6), anteverted nares (Families 4, 6), and bulbous nasal tip (Families 3, 5, 6).

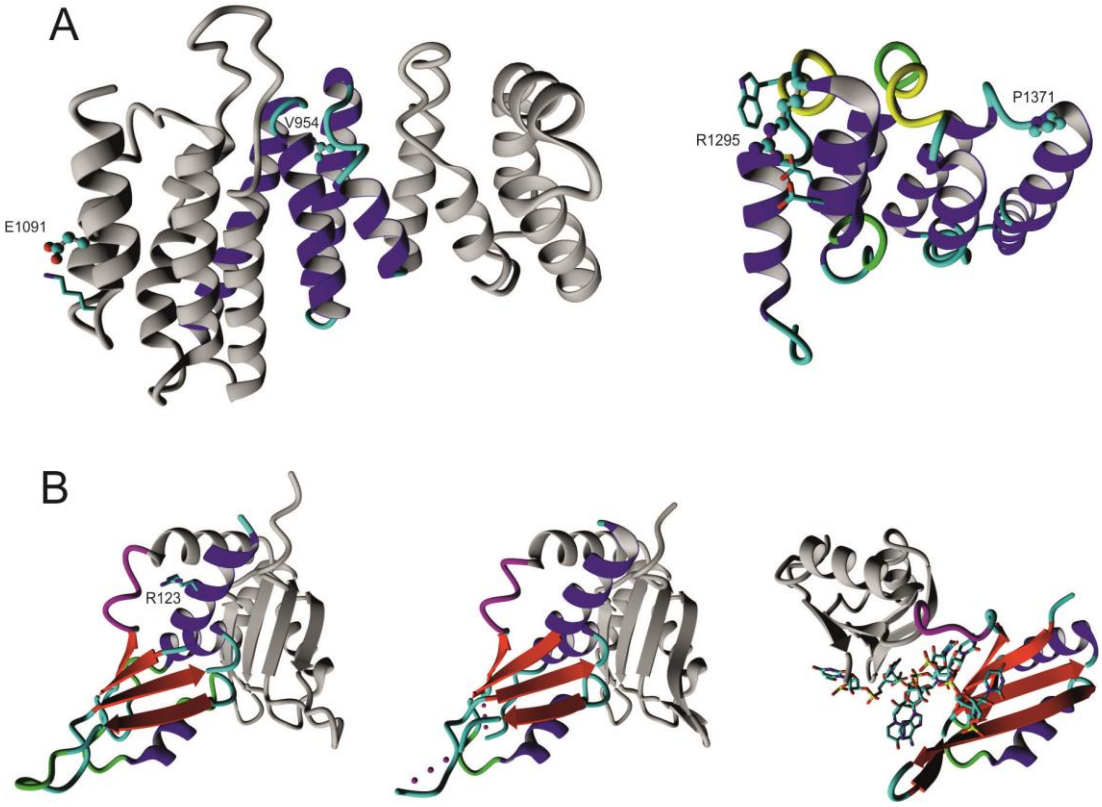
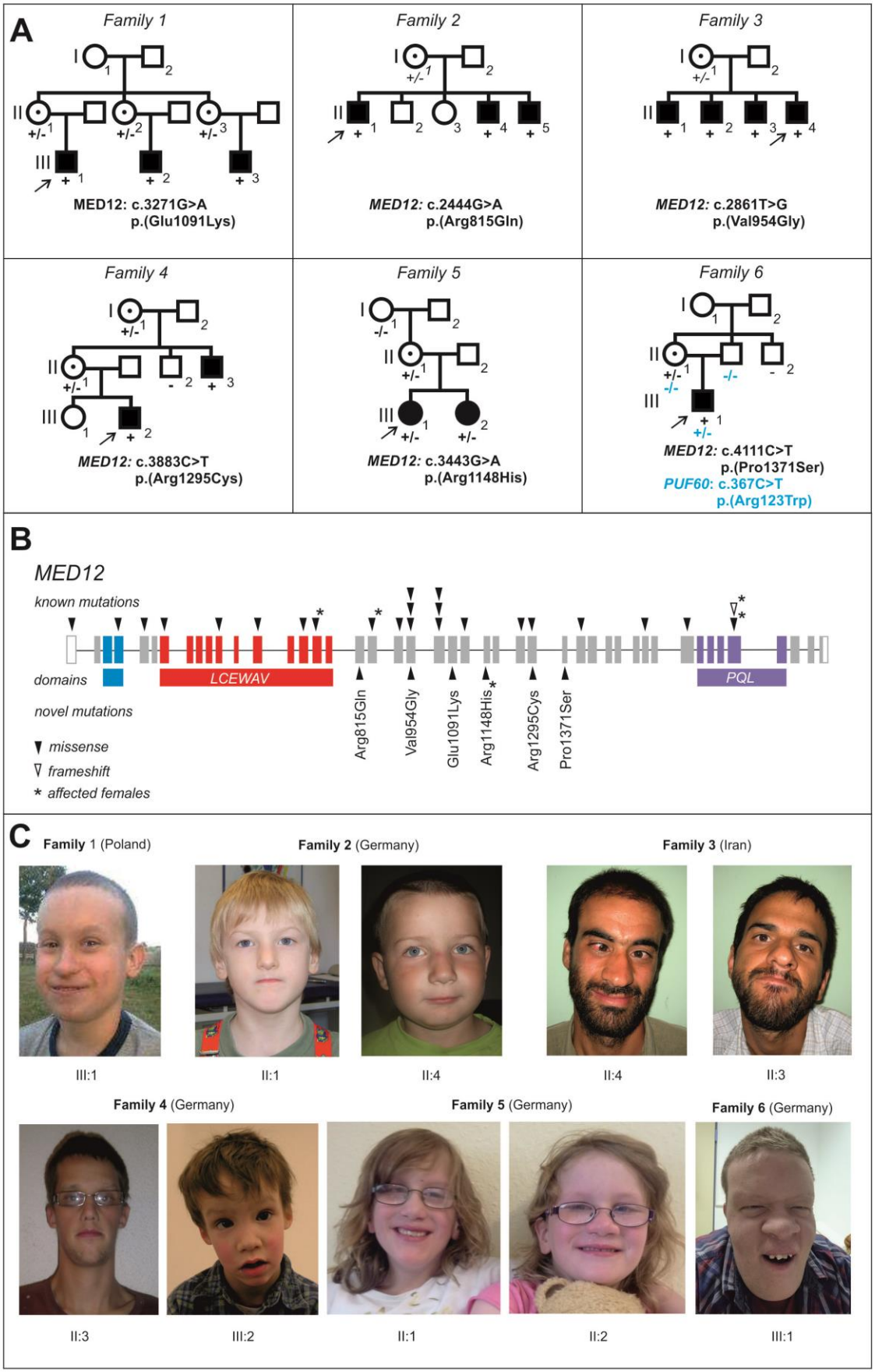


Fig. 2

A Representative models (Yasara Structure) of the MED12 segments 826Ala-1096Leu and 1278Tyr-1395Glu which harbor variants identified in the families, obtained based on the structures of homologous proteins. Glu1091 (E1091) forms a salt bridge with the proximal Lys1086, so the missense variants p.(Glu1091Lys) (E1091K), which replaces this interaction by an unfavorable Lys-Lys one (cation-cation), decreases protein stability. The amino acid substitution Val954Gly (V954G) decreases the protein stability distorting the hydrophobic core of the four helix bundle (blue). Arg1295 (R1295) forms a salt bridge with Glu1302 (E1302) and/or Asp1303 (D1303) and displays pi-pi interactions with Trp1291 (W1291), thus stabilizing the relative orientation of two proximal helices. In this view, the p.(Arg1295Cys) substitution will destroy all these interactions. It should be noted that three cysteine residues are proximal to Arg1295 (Cys1287, Cys1296 and Cys1301), so the Arg1295Cys replacement, which introduces another cysteine residue, enables formation of alternative patterns of Cys-Cys bridges, all of which will interfere with the native one, thus affecting the protein folding. Pro1371Ser (P1371S) will increase the local flexibility of the short loop, thereby enabling reorientation of the flanking helices, and thus decreasing protein stability.



B The analysis (Yasara Structure) of the PUF60 p.(Arg123Trp/Gly) amino acid substitutions. Arg123 (R123) is located just before the RRM1/RRM2 motifs of PUF60, in the proximity of the linker ^{205}ProSerAsnIleGlyGln^{210} (magenta) that separates RRM domains (left). Arg123Gly (R123G) replacement only minutely affects the protein structure (middle). The change of the polar arginine to the highly hydrophobic tryptophan restricts the conformational flexibility of this linker (right), and the relative orientation of both RRM domains changes substantially upon RNA binding, thus most likely affecting the affinity of PUF60 to RNA.

Feature category	*Feature*	*FG Syndrome/ Opitz-Kaveggia*	*Lujan-Fryns Syndrome*	*Ohdo Syndrome*	*Family 1*	*Family 2*	*Family 3*	*Family 4*	*Family 5*	*Family 6*
General	Intellectual disability	+	+	+	+	+	+	+	+	+
	Developmental delay	+	+	+	+	+	+	+	+	+
	Poor/absent speech			+			+			+
	Speech delay/speech disturbances	+		+	+	+			+	+
	Hypernasal speech		+		+					+
	Short stature	+			+	+				+
	Tall, thin body habitus (marfanoid features)		+					+		
Central nervous system	Hypotonia	+	+	+			+	+		+
	Spasticity with joint contractures	+			+					
	Seizures and EEG abnormalities	+	+	+				+		+
	Corpus callosum agenesis	+	+	+				+		+
	Mega cisterna magna					+				
Craniofacial	Macrocephaly	+	+		+/-		+		+	+
	Microcephaly			+		+				
	Dolichocephaly	+	+							
	Brachycephaly					+				+
	Long, narrow face	+	+			+	+	+		+
	Triangular face			+						
	Coarse face			+						+
	Prominent forehead	+	+	+			+	+	+	+
	Malar flattening		+		+					
	Prominent nasal bridge		+			+		+	+	
	Short philtrum		+						+	
	Thick alae nasi			+	+					
	Small nares								+	
	Anteverted nares							+		+
	Bulbous nasal tip						+		+	+
	Small ears	+		+	+			+	+	+
	Large ears					+				
	Open mouth	+	+							+
	Narrow lips			+	+				+	
	High, arched palate	+	+	+	+			+		
	Cleft Lip/Palate			+		+				
	Dental abnormalities		+	+	+	+		+		+
	Maxillary hypoplasia			+						
	Micrognathia	+	+	+		+				
	Prognathism				+		+		+	+
	Facial ticks and grimacing					+				
	Hypertelorism	+		+				+	+	+

	Hypotelorism		+		+					
	Blepharophimosis			+						
	Ptosis		+	+						+
	Puffy eyelids	+								
	Epicanthal Folds			+					+	+
	Downslanted palpebral fissures	+	+	+						+
	Upslanted palpebral fissures							+		
	Deep set eyes				+				+	
	Sparse eyebrows		+					+		+
Ophtalmologic	Nystagmus	+								
	Strabismus	+	+	+			+	+	+	
Auditory	Hearing loss			+						+
Musculosceletal	Skeletal anomaly	+		+				+		
	Wide neck				+					
	Broad thumbs and halluces	+	+							
	Long hands		+		+					
	Camptodactyly			+						+
	Single palmar crease							+		+
	Hyperextensible joints		+	+		+				
Cardiopulmonary	Cardiac anomaly	+	+	+				+		+
Gastrointestinal	Feeding problems	+		+	+					
	Gastroesophageal reflux	+								
	Hirschsprung disease			+						+
	Constipation	+		+	+					+
	Anal anomaly	+								+
Genitourinary	Cryptorchidism	+		+		+		+		+
	Genital hypoplasia			+						+
	Inguinal hernia	+			+			+		
	Urinary incontinence					+				
Behaviour	Hyperactive, friendly, affable	+				+	+			+
	Hyperactive, aggressive, shy		+	+	+	+				
	Psychosis		+		+					

Table 1. Phenotypic comparison between known MED12 related syndromes and the families presented in the study. Clinical symptoms specific for a given syndrome are coloured appropriately (FG, red; L-F, green; Ohdo, blue). Symptoms not described previously that expand the phenotypic spectrum of the *MED12* variants are marked in bold. "+" in family columns denotes that at least one patient per family was diagnosed with the given feature, "blank" denotes that none of the family members presented the given feature.