De novo and biallelic *DEAF1* variants cause a phenotypic spectrum

Authors: Maria J. Nabais Sá MD PhD[1,‡], Philip J. Jensik PhD[2,‡], Stacey R. McGee MS[2], Michael J. Parker MD[3], Nayana Lahiri MD[4], Evan P. McNeil MS[5], Hester Y. Kroes MD PhD[6], Randi J. Hagerman MD[7,8], Rachel E. Harrison MD[9], Tara Montgomery MD[10], Miranda Splitt MD[10], Elizabeth E. Palmer MBBS BA FRACP[11,12], Rani K. Sachdev MD[11,12], Heather C. Mefford MD PhD[13], Abbey A. Scott BS MS[14], Julian A. Martinez-Agosto MD[15,16], Rüdiger Lorenz MD[17], Naama Orenstein MD[18,19], Jonathan N. Berg MD[20,21], Jeanne Amiel MD[22], Delphine Heron MD[23], Boris Keren MD PhD[23], Jan-Maarten Cobben MD[24,25], Leonie A. Menke MD PhD[24], Elysa J. Marco MD[26], John M. Graham Jr. MD ScD [27], Tyler Mark Pierson MD PhD[28], Ehsan Ghayoor Karimiani MD PhD[29], Reza Maroofian PhD[29], M. Chiara Manzini PhD[30], Edmund S. Cauley MSc[30], Roberto Colombo MD[31,32], Sylvie Odent MD[33], Christele Dubourg Pharm D PhD [34], Chanika Phornphutkul MD[35], Arjan P.M. de Brouwer PhD[1], Bert B. A. de Vries MD PhD[1,∞], Anneke T. Vulto-vanSilfhout MD PhD[1]

‡ Equal Contribution Status.
∞ This is the corresponding author.

1 Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands;
2 Department of Physiology, Southern Illinois University School of Medicine, Carbondale, IL, USA;
3 Sheffield Clinical Genetics Service, OPD2 Northern General Hospital, Sheffield, UK;
4 Department of Clinical Genetics, St George's University Hospitals NHS Foundation Trust & St George's, University of London, London, UK;
5 Dartmouth Geisel School of Medicine, Hanover, New Hampshire, USA;
6 Department of Genetics, University Medical Center Utrecht, Utrecht, The Netherlands;
7 Medical Investigation of Neurodevelopmental Disorders (MIND) Institute, University of California Davis School of Medicine, Sacramento, California, USA;
8 Department of Pediatrics, University of California Davis Medical Center, Sacramento, California, USA;
9 Department of Clinical Genetics, Nottingham University Hospitals NHS Trust, Nottingham, UK;
10 Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK;
11 Sydney Children's Hospital, Randwick, NSW, Australia;

12 School of Women's and Children's Health, UNSW Medicine, The University of New South Wales, Sydney, NSW, Australia;
13 Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, Washington, USA;
14 Division of Genetic Medicine, Seattle Children's Hospital, Seattle, Washington, USA;
15 Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA;
16 Division of Medical Genetics, Department of Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA;
17 Ludwig-Konrad-Str. 14, 34537 Bad Wildungen, Germany;
18 Pediatric Genetics Clinic, Schneider Children's Medical Center of Israel, Petach Tikva, Israel;
19 Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel;
20 Department of Clinical Genetics, Ninewells Hospital and Medical School, Dundee, Angus, UK;
21 Clinical Genetics, University of Dundee, Dundee, Angus, UK;
22 Département de Génétique, Hôpital Necker-Enfants Malades, Assistance Publique, INSERM UMR 1163, Institut Imagine, Paris, France;.
23 Département de Génétique, Hôpital Pitié-Salpêtrière, Assistance publique - Hôpitaux de Paris, Paris, France;
24 Department of Pediatrics, Amsterdam University Medical Centers, Amsterdam, The Netherlands;
25 North West Thames Genetics NHS, Northwick Park Hospital, London, UK;
26 UCSF Pediatric Brain Center, Sensory, Neurodevelopment & Autism Program (SNAP), University of California, San Francisco, CA, USA;
27 Division of Clinical Genetics and Dysmorphology, Department of Pediatrics, Cedars-Sinai Medical Center, Los Angeles, CA, USA;
28 Department of Pediatrics, Department of Neurology, and the Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA;
29 Genetics Research Centre, Molecular and Clinical Sciences Institute, St George's, University of London, London, UK;
30 GW Institute for Neuroscience, Department of Pharmacology and Physiology, The George Washington University School of Medicine and Health Sciences, Washington, DC, USA;
31 Faculty of Medicine "Agostino Gemelli" - Catholic University of the Sacred Heart, Rome, Italy;
32 Center for the Study of Rare Inherited Diseases (CeSMER), Niguarda Ca' Granda Metropolitan Hospital, Milan, Italy;
33 Service de Génétique Clinique, CLAD-Ouest CHU Rennes, Univ Rennes, CNRS 6290 Institut de Génétique et Développement de Rennes (IGDR), Rennes, France;
34 Univ Rennes, CHU Rennes, CNRS, IGDR, UMR 6290, Rennes, France;
35 Division of Human Genetics, Department of Pediatrics, Hasbro Children's Hospital, Warren Alpert Medical School of Brown University, Providence, RI, USA.

Conflict of interest: The authors declare no conflict of interest.

Corresponding author: Bert B. A. de Vries, Department of Human Genetics (route 836), Radboud University Medical Center, P.O. box 9101, 6500 HB, Nijmegen, The Netherlands; T: +31 (0)24 36 53678; E-mail: Bert.deVries@radboudumc.nl.

Abstract

Purpose: To investigate the effect of different *DEAF1* variants on the phenotype of patients with autosomal dominant and recessive inheritance patterns and on DEAF1 activity *in vitro*. **Methods:** We assembled a cohort of 23 patients with *de novo* and biallelic *DEAF1* variants, described the genotype-phenotype correlation and investigated the differential effect of *de novo* and recessive variants on transcription assays using DEAF1 and Eif4g3 promoter luciferase constructs. **Results:** The proportion of the most prevalent phenotypic features, including intellectual disability, speech delay, motor delay, autism, sleep disturbances, and a high pain threshold, were not significantly different in patients with biallelic and pathogenic *de novo DEAF1* variants. However, microcephaly was exclusively observed in patients with recessive variants ($p<0.0001$). **Conclusion:** We propose that different variants in the *DEAF1* gene result in a phenotypic spectrum centered around neurodevelopmental delay. While a pathogenic *de novo* dominant variant would also incapacitate the product of the wild-type allele and result in a dominant-negative effect, a combination of two recessive variants would result in a partial loss-of-function. Since the clinical picture can be non-specific, detailed phenotype information, segregation, and functional analysis are fundamental to determine the pathogenicity of novel variants and to improve the care of these patients.

Keywords: *DEAF1*, neurodevelopmental disorder, intellectual disability, genotype, phenotype

Introduction

DEAF1 (MIM*602635; NM_021008.3) encodes the deformed epidermal autoregulatory factor-1 homolog (DEAF1), a transcription factor that is highly expressed in the central nervous system, in particular during early embryogenesis.[1] DEAF1 regulates the expression of various genes[2] as both a transcriptional activator and repressor.[3-6] It contains several functional domains including a centrally located SAND (Sp-100, AIRE, NucP41/75, and DEAF1) domain, a zinc finger motif (ZnF), a nuclear localization signal (NLS), a nuclear export signal (NES), and an MYND (myeloid translocation protein 8, Nervy, and DEAF1) domain (Figure 1).[6,7] The region encompassing the SAND domain constitutes a DNA binding domain, which binds to TTCG motifs.[6] The SAND and MYND domains are also involved in protein-protein interactions.[6-9] The presence of a NLS and NES indicates that DEAF-1 may be regulated by nuclear/cytoplasmic shuttling.[7] Biallelic disruption of *Deaf1* in mice results in neural-tube defects[10], and biallelic *Drosophila Deaf1* loss-of-function mutants show early embryonic arrest[11].

Pathogenic variants in the *DEAF1* gene have been reported to lead to two clinically distinct intellectual disability (ID) syndromes: autosomal dominant mental retardation-24 (MRD24; MIM#615828) caused by *de novo* variants,[1,12-17] and the recessively inherited dyskinesia, seizures, and intellectual developmental disorder syndrome (DYSEIDD; MIM#617171)[16,18-20]. These two syndromes are collectively described as *DEAF1*-associated neurodevelopmental disorders (DAND). To date, nine different *de novo* pathogenic variants have been described in ten individuals, who manifested moderate to severe ID with severely affected expressive speech and

mild motor delay.[1,12-17] Epilepsy was described in half of the individuals.[16] Behavioral problems consisted of hyperactive, compulsive and/or aggressive behavior, fascination with water, and striking mood swings.[1,14-16] Other characteristics were autism, recurrent infections, a high pain threshold, and an abnormal walking pattern.[1,14-16] Body measurements were normal, facial dysmorphisms were only mild, and no other major congenital anomalies were observed.[1] All but one of the *de novo* pathogenic *DEAF1* variants occurred in the SAND domain, including seven missense variants and one splice site variant.[1,12-17] One in-frame deletion of three base pairs was reported in the NLS domain.[16] These *de novo* pathogenic *DEAF1* variants impair the *DEAF1* transcriptional activity, DNA-binding, and/or alter subcellular localization.[1,16] Since heterozygous *DEAF1* deletions[21] do not cause DAND, *de novo* pathogenic variants have been proposed to lead to a dominant-negative effect.[1]

Two kindreds with patients with homozygous DYSEIDD pathogenic *DEAF1* variants were previously reported, a missense change and a non-canonical splice site pathogenic variant.[16,18-20] In total, seven patients homozygous for either of these variants presented with ID, microcephaly, and hypotonia, were described.[16,18-20] Seizures and white matter abnormalities resulting in dyskinesia were reported in four out of the six patients for which these features were investigated.[18-20] The missense variant was present in the SAND domain as well, whereas the splice site variant resulted in exon skipping and reduced the normal full-length mRNA copy number in the patients to 5% of the wild-type level,[19] suggesting that the recessive phenotype resulted from a partial loss-of-function of *DEAF1*.

In order to investigate the effect of different *DEAF1* variants on the phenotype of patients with autosomal dominant and recessive inheritance patterns, we have assembled a cohort of 23 patients not reported before with *de novo* and biallelic *DEAF1* variants. We describe the genotype-phenotype correlation and investigate the differential effect of *de novo* and recessive variants on DEAF1 activity *in vitro*.

Patients & Methods

Identification of individuals

Patients were referred by physicians working in intellectual disability support services and by clinical geneticists. Written informed consent was obtained for all individuals involved. This study adhered to the World Health Association Declaration of Helsinki (2013) and was approved by the institutional review board Commissie Mensgebonden Onderzoek Regio Arnhem-Nijmegen. Genomic DNA was isolated from peripheral blood samples following standard procedures.[22] Exome enrichment, high-throughput sequencing, and subsequent *de novo* analysis was performed by established procedures.[23] Copy number variant analysis was performed on exome data by using CoNIFER[1,24] as described before.[25]

Plasmid constructs

DEAF1 mammalian expression and *DEAF1* and *Eif4g3* promoter luciferase plasmids have been previously described.[1] Site-directed mutagenesis was used to introduce the specific amino acid substitutions into the DEAF1 expression plasmids, derived from the human DEAF1 cDNA (GenBank accession number AF049459). Variants resulting in the indicated human amino acid substitutions were generated by PCR as previously described.[1]

Transcription assays

Luciferase assays using *DEAF1* and *Eif4g3* promoter luciferase constructs have been previously described.[1] Briefly, HEK293T cells in 24-well plates were transfected

with 125 ng pcDNA3 (control) or DEAF1 (WT or DEAF1 variants) expression plasmids with 375 ng promoter-luciferase and 1.25 ng RSV-Renilla luciferase constructs using the calcium phosphate technique for 18 h. Luciferase assays were performed 24 hours later using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI USA).

DEAF1 EBV-LCL RNA analysis

Epstein Barr virus-transformed lymphoblastoid cell line (EBV-LCL) cells ($8x10^6$ cells) were grown in 20 mL RPMI supplemented with 15% fetal bovine serum, penicillin, and streptomycin. Twenty-four hours later 10 mL was removed to a new flask and cells were treated with vehicle (water) or cycloheximide (50 ug/mL) for 6 hours. RNA was isolated from EBV-LCL cells using Trizol reagent and 1.0 ug of RNA was reverse transcribed. PCR was performed using cDNA and primers to amplify specific regions of DEAF1 cDNA for downstream restriction digests or subcloning and DNA sequencing. Quantitative PCR was performed using DEAF1 (1617dup) and GAPDH primers[26] and data were normalized using the $2^{(-\Delta\Delta Ct)}$ method.[27]

Immunoprecipitation and Western Blot

EBV-LCL were lysed in lysis buffer containing 150 mM sodium chloride, 50 mM Tris (pH=7.5), 1.0% TritonX-100, 1.0 mM EDTA, 1.0 mM sodium fluoride, 0.2 mM sodium orthovanadate, aprotinin (10 μg/mL), leupeptin (10 μg/mL) and pepstatin (10 μg/mL) and lysates were incubated with rabbit DEAF1 antibodies[8] or preimmune serum bound to protein G magnetic beads overnight at 4°C. Beads were washed and

proteins were eluted in Laemmli sample buffer. Transfected HEK293t cells were lysed in lysis buffer. Proteins were separated on SDS-PAGE gels and transferred to polyvinylidene fluoride membranes. Immunoblot analysis was performed with a different rabbit anti-DEAF1 antibody for EBV-LCL immunoprecipitation experiments or rabbit anti-DEAF1 and mouse anti-bACTIN (Abcam, Cambridge, MA USA) for transfected HEK293t experiments.

Genotype-phenotype correlations

Fisher's Exact Test, using a two-sided p-value, was carried out to compare proportions of clinical features between groups of patients with *de novo* and biallelic *DEAF1* variants. After conducting Bonferroni correction, a p-value of <0.0025 was considered significant.

Results

De novo *variants in the SAND domain of* DEAF1 *impair transcriptional activity*

We identified 14 novel and two previously published *de novo DEAF1* variants in a total of 18 patients by exome sequencing (n=16), genome sequencing (n=1), autism gene panel analysis (n=1) performed in 17 medical centres from all over the world (Figure 1; Table S1). These variants included 14 missense variants, one *in-frame* deletion that resulted in a single amino acid change, and one splice site variant, c.664+1G>T. To determine if the c.664+1G>T variant resulted in alternatively spliced DEAF1 mRNA transcripts, PCR was used to amplify regions containing exons 2-5 using Epstein Barr virus-transformed lymphoblastoid cell line (EBV-LCL) mRNA from individual AD/8. DNA sequencing indicated the variant causes skipping of exon 4, which results in an in-frame deletion of 49 amino acid residues in the SAND domain p.(Pro174_Gly222del) (Figure 2). Compared to control EBV-LCL, a truncated DEAF1 protein, as well as full length DEAF1, were observed in patient AD/8 EBV-LCL lysate and corresponded to a similar molecular weight protein using lysates from cells transfected with WT or p.(Pro174_Gly222del) plasmid DNA. The novel variant p.(Ala276Pro) was found in two unrelated individuals from this study. The previously reported *de novo* variants p.(Gly212Ser) and p.(Gln264Pro)[1,13,16] were identified in our cohort in two and one families, respectively. The p.(Gly212Ser) variant was identified in two out of the 247 reads in peripheral blood cells of the patient's unaffected father. The variant was not detected by Sanger sequencing, suggesting that he is likely a mosaic for this variant.

Fifteen of the sixteen *de novo* variants are located in the SAND or adjacent ZnF domain; the p.(Pro293Leu) variant is between the ZnF and NLS (Figure 1c). All missense variants were predicted by SIFT[28] and Polyphen-2[29] to be disease-causing and had a CADD score >20.0.[30] None of the novel *de novo* variants were found in gnomAD (Genome Aggregation Database).[31] Multispecies alignment showed that all mutated amino acids are highly conserved across species (Figure S1).

DEAF1 binds and represses its own promoter activity.[32] *DEAF1* promoter transcriptional repression activity was analyzed for *de novo* variants. All variants lost transcriptional repression activity as compared to wild type DEAF1, although p.(Pro293Leu) retained some repression activity (Figure 3A). DEAF1 also increases the transcriptional activity of the mouse *Eif4g3* promoter[5] and *de novo DEAF1* variants have previously been shown to suppress activation of this promoter.[1] In 12 out of our 14 variants, an approximate 10-fold suppression of transcriptional activity relative to basal expression levels was demonstrated. No activation or reduction in promoter activity was observed for p.(Ser236Gly) relative to basal activity, and p.(Pro293Leu) increased promoter activity (Figure 3B).

Biallelic variants in DEAF1 *result in no significant effect in transcriptional assays*

Six novel biallelic *DEAF1* variants were found in compound heterozygosity in three unrelated individuals and one novel homozygous variant was identified in two siblings of a consanguineous family by using exome sequencing (Figure 1d; Table S2). These variants were scattered throughout *DEAF1*. In two families, compound

heterozygous variants, p.(Trp234*) and p.(Glu239Gly), or a homozygous variant, p.(Arg224Gln), were found in the SAND domain. The compound heterozygous missense variant p.(Glu239Gly) is not reported in gnomAD and is predicted by SIFT to be deleterious, by Polyphen-2 to be probably damaging, and had a CADD score 25. The homozygous substitution p.(Arg224Gln) located in the SAND domain is present only once heterozygously in gnomAD (frequency of 4.065e-6), was predicted by SIFT to be tolerated, by Polyphen-2 to be disease-causing, and had a CADD score 31. In the other two families, a single base pair duplication or deletion resulting in frameshift near to the 5' of *DEAF1*, p.(Arg44Glyfs*25), or in the middle of the *DEAF1* gene, p.(Asp369Alafs*51), both located outside a domain, occurred in combination with a variant in the MYND domain, respectively an *in-frame* deletion, p.(Phe527del) or an insertion resulting in a frameshift, p.(Cys540Metfs*18). Three of these variants, p.(Arg44Glyfs*25), p.(Trp234*), and p.(Asp369Alafs*51) are expected to lead to nonsense-mediated decay (NMD) of their respective transcripts.

DEAF1 mRNA expression analysis was performed using EBV-LCL from patient AR/1, who carried c.1617dup, p.(Cys540Metfs*18), and c.1104_1105dup, p.(Asp369Alafs*51) in compound heterozygosity, and from parents carrying either the c.1617dup or c.1104_1105dup. To differentiate the expression of DEAF1 transcripts from each allele, two different PCR amplicons were generated that contain either c.1617dup or c.1104_1105dup variant from patient and parent EBV-LCL treated with vehicle or cycloheximide (CHX), an inhibitor of NMD. The c.1617dup eliminates an endogenous *Nde*I site and the c.1104_1105dup generates a new *Hha*I site. Restriction endonuclease digestion analysis indicated that both

c.1617dup or c.1104_1105dup containing mRNA transcripts were present in patient and parent cDNA. The c.1104_1105dup transcript seemed to be present in a lower amount in cells not treated with CHX as compared to those treated with CHX. As CHX is an inhibitor of NMD, this suggests that this transcript is subject to NMD (Figure S2). DEAF1 immunoprecipitaions were performed on EBV-LCL lysates from patient and parents. Compared to parent EBV-LCL, no full length or p.(Cys540Metfs*18) ~80 kDa DEAF1 protein was observed in patient lysate. A faint truncated ~60kDa band consistent with the p.(Asp369Alafs*51) DEAF1 protein was observed in patient and parent c.1104_1105dup. Quantitative PCR (qPCR) showed that *DEAF1* mRNA expression in untreated EBV-LCL of the patient AR/1 (c.1104_1105dup/c.1617dup variants) was 30% of normal EBV-LCL expression levels ($p<0.05$), Figure S3). CHX treatment significantly increased *DEAF1* mRNA expression relative to vehicle treated cells in c.1104_1105dup/c.1617dup patient as well as the parent carrying the c.1104_1105dup supporting that the c.1104_1105dup containing transcript is targeted by NMD

The effects of the biallelic variants on DEAF1 transcriptional activity, protein expression, and cellular localization were also determined. Compared to WT DEAF1, no change in transcriptional repression activity was observed for the biallelic variants, except for p.(Trp234*) (Figure 3C). No significant change in transcriptional activation of the *Eif4g3* promoter was observed. (Figure 3D). Interestingly, p.(Arg44Glyfs*25), which should result in a severely truncated DEAF1 protein that lacks most of the DEAF1 protein domains including the SAND domain, retained transcriptional activity. Western blots were performed using lysates from HEK293t

transfected with WT or DEAF1 biallelic variants. A faint ~70 kDa protein was observed for p.(Arg44Glyfs*25) (Figure S4). This might be the result of the use of a translational start at a downstream in-frame AUG start codon located at codon 69. The p.(Arg44Glyfs*25) variant results in a stop codon 5' adjacent to methionine 69 of *DEAF1*. The p.(Trp234*) and p.(Asp369Alafs*51) variants resulted in truncated DEAF1 proteins approximately 30 and 60 kDa in size, respectively. The p.(Cys540Metfs*18) variant resulted in a slightly lower molecular weight protein. As expected p.(Arg224Gln), p.(Glu239Gly), and p.(Phe527del) variants had the same apparent molecular weight as WT DEAF1. Immunofluorescent staining was used to determine the subcellular localization of the biallelic *DEAF1* variants in transfected cells. Unlike WT DEAF1, which localized to the nucleus, p.(Trp234*) was found throughout the cell (Figure S5). The other biallelic variants localized to the nucleus.

De novo DEAF1 variants result in a non-specific phenotype comprising ID, speech delay, motor delay, sleep disturbance, autism, and a high pain threshold

Fourteen male and three female individuals with a *de novo DEAF1* variant (median age: 7.5 years; age range: 2.5 years – 38 years), were clinically characterized (Table 1; Table S3). All were born from healthy unrelated non-consanguineous families. This cohort consisted of eleven children, six of which were preschool children, three were adolescents and three were adults. One individual was excluded from genotype-phenotype analysis, as he had the variant p.(Pro293Leu), which was not located in the SAND domain and did not result in suppression, but rather an increase of Eif4g3 promoter activity.

Developmental delay (DD)/ID was present in all patients (17/17; 100%) and was typically moderate-severe or severe DD/ID (Table S4). Motor delay, present in most of the patients (12/17; 71%), was usually mild. Language development was severely delayed in all patients, except for one patient with the variant p.(Leu214Val) (16/17; 94%). Seven patients had absent speech and nine had limited speech. Of note, nonverbal patients were said to have used a few words before losing the ability to produce speech. Developmental regression was described in half of the individuals, mostly corresponding to loss of the ability of communicating verbally. Although in most individuals expressive language was severely impaired, receptive language was perceived, particularly by caregivers, as better than expressive. Furthermore, patients (15/15; 100%) were able to communicate simple needs nonverbally, for instance taking the caregivers' hand to point or lead to the item(s) of interest, using gestures and sign language, producing sounds accompanied (or not) by gestures (e.g. to express emotions, as annoyance), or using picture exchange communication system, picture prompts, a tablet, or a touch-based speech-generating device.

Autism spectrum disorder occurred in all patients (16/16; 100%). In the majority, poor eye contact (12/16; 75%), fascinations (15/17; 88%), particularly a compulsive interest in water, and sudden mood swings (16/17; 94%) were concomitantly observed. Mood swings were characterized by exaggeration of emotional affects, with paroxysmal bursts of laughter and context-inappropriate happy disposition alternating with aggressive behavior or depressive feelings. Aggressive behavior (12/16; 75%) manifested both as hetero-aggression and auto-

aggression (e.g. self-mutilation, head-banging, self-biting). Other behavioral problems (13/15; 87%) included pica, impulsivity/hyperactive behavior, and obsessive traits.

Epilepsy was also observed in the majority of patients (10/14; 71%). Median age at diagnosis was three years, although age of onset varied from two days old to 16 years old (Table S5). Seizures were usually generalized and frequently difficult to treat or intractable (9/10; 90%). Anti-epileptic drugs (AED) were able to reduce seizures in five patients, while in four other patients seizures were refractory to treatment. Of note, seizures were noticed to adversely influence speech.

Sleep dysfunction was frequent and severe (15/17; 88%). It consisted of difficulty falling asleep (sleep onset insomnia) and awaking frequently (maintenance insomnia). Patients were reported to be very disruptive when waking during the night. In some patients, sleeping problems were treated with some benefit using medication, such as melatonin.

Other neurologic abnormalities included a remarkably high pain threshold (14/16; 88%), hypotonia (10/16; 63%) and gait difficulties, comprising of gait ataxia (4/16; 25%) or an ataxic looking gait (12/17; 71%), as broad-based and imbalanced gait, or a tip-toe gait. Brain abnormalities as indicated by brain MRI were present in less than a third of the patients (4/13; 31%), but no common pattern of brain malformations could be distinguished. Movement disorder, in particular dystonia, was only reported in one individual.

Non-neurodevelopmental issues included recurrent infections that occurred in more than half of patients (10/16; 63%), mostly in infancy or childhood. They

included predominantly chronic ear infections, sinusitis, or upper respiratory tract infections. In addition, gastrointestinal abnormalities were common (12/16; 75%) and included feeding difficulties, gastroesophageal reflux disease, constipation, and diarrhea. Craniofacial dysmorphisms were mild (Figure 1e) and macrocephaly was observed in approximately one quarter of individuals (4/17; 24%). Growth was usually within the norm for age and gender.

Biallelic DEAF1 variants result in microcephaly

Microcephaly (3/5; 60%) was exclusively observed in patients with recessive variants. All patients with biallelic *DEAF1* variants manifested ID and developmental regression, except for one (Tables S6 and S7). All parents were healthy. Remarkably, although the patient with variants p.(Asp369Alafs*51) and p.(Cys540Metfs*18) in compound heterozygosity never learned to walk and had absent speech, similarly to other patients with biallelic *DEAF1* variants, he completed various exams from the first grade of secondary school, suggesting he has a normal cognitive function, although this was not formally tested. Additionally, he used non-verbal communication aids (supportive typing) that revealed good receptive language. Behavioral problems occurred in all patients and ranged from poor eye contact (4/5; 80%) to autism (1/4; 25%) and aggressive behavior (2/5; 40%). Frequent neurologic abnormalities included epilepsy (4/5; 80%) (Tables S5), sleep disturbance (5/5; 100%), and brain abnormalities (4/5; 80%), such as ventriculomegaly with or without hydrocephalus, corpus callosum hypoplasia or white matter abnormalities. Movement disorder (2/4; 50%) included dystonia and

stereotypies. Common extra-neurologic features included feeding difficulties (3/5; 60%) and constipation (5/5; 100%). Lastly, facial dysmorphisms were minor and non-specific (Figure 1e).

De novo and biallelic variants result in a similar phenotype except for microcephaly

The proportion of the most prevalent phenotypic features, including ID, speech delay, motor delay, autism, sleep disturbances, and a high pain threshold, was not significantly different in patients with biallelic and pathogenic *de novo* *DEAF1* variants, when considering all patients included in this study and previously described individuals (Table 1). Interfamilial phenotypic variability was observed among patients with the same *de novo* variant (*e.g.* some patients developed seizures, while others did not). Microcephaly, however, was observed exclusively in patients with recessive variants ($p<0.0001$).

Discussion

In this study, we expanded the clinical consequences of variants implicated in *DEAF1*-associated neurodevelopmental disorders (DAND) by describing the phenotype of 17 patients with a *de novo* pathogenic variant and five patients with biallelic variants in the *DEAF1* gene. We determined that the *de novo* variants impaired DEAF1 transcriptional repression activity and fifteen of the seventeen variants reversed DEAF1-mediated transcriptional activation at the *Eif4g3* promoter. Apart from the p(Trp234*) variant, the other six novel recessive *DEAF1* variants demonstrated no change in transcriptional activity.

We identified missense variants in the SAND domain, both *de novo* and recessive. Only *de novo* variants, however, altered *DEAF1* and *Eif4g3* promoter activity. This is in line with previous data implying a dominant-negative effect of these variants.[1] So far, no *de novo* nonsense or frameshift variants resulting in disease have been identified. *De novo* variants clustered in very close proximity within the SAND domain. Moreover, variants have been found recurrently affecting the same residues, namely p.Gly212, p.Leu214, p.Lys216, p.Ile228, p.Arg254, p.Gln264 and p.Ala276.[1,12-20] Additionally, the canonical donor splice site 3' of c.664 was also found to be repeatedly affected.[17] We show in our study that c.664+1G>T causes skipping of exon 4, producing an in-frame deletion in the SAND domain that results in abnormal transcriptional activity mimicking the effect of other missense variants on DEAF1 protein activity.

Missense biallelic variants in the SAND domain do not impair the function of this domain as evidenced by the normal transcriptional activity of the mutant

protein in our assays. This is especially remarkable for the novel homozygous variant p.(Arg224Gln), as the previously reported *de novo* p.(Arg224Trp) variant was shown to result in loss of transcriptional repression of the *DEAF1* promoter, and was unable to activate the *Eif4g3* promoter.[1] We hypothesize that the p.(Arg224Gln) variant affects normal protein folding of DEAF1 because arginine, an amino acid with an electrically charged side chain, is replaced by an amino acid with a polar uncharged chain. This is also the case for the recessive p.(Arg226Trp) variant, also located in the SAND domain.[16,18,20] Similarly, the compound heterozygous variant p.(Glu239Gly), which results in a protein with an apparently normal length, may have also an effect on the normal protein folding, due to the substitution of the electrically charged side chain of a glutamate for a glycine with a small non-polar side chain. It is possible the DEAF1 variants could influence other aspects of DEAF1 function or activity that may not be detected by the cell-based functional assays used in this study. Thus, we predict that these variants constitute a classical, but incomplete loss-of-function (hypomorph).

Loss-of-function recessive variants in *DEAF1* consist of p.(Arg44Glyfs*25), p.(Trp234*), p.(Asp369Alafs*51), and p.(Cys540Metfs*18). Only p.(Trp234*) affected transcriptional repression activity. This variant likely results in a truncated protein lacking most of the SAND domain. Our results suggest that the DEAF1 transcript containing the p.(Arg44Glyfs*25) variant can be transcribed using an alternative start codon resulting in a N-terminal deletion of 69 amino acid residues, and thus in an intact SAND domain. Accordingly, these data indicate the MYND domain is not essential for DEAF1 transcriptional repression or activation activity

using the *in vitro* assays in this study. Taken together, the *de novo* SAND domain variants likely have a dominant-negative effect (antimorph), whereas the biallelic *DEAF1* variants lead to a partial loss-of-function (hypomorph) through either a reduction in function with normal protein levels (missense variants) or reduced protein levels due to NMD (nonsense and frameshift variants). This hypothesis also explains why heterozygous deletions of *DEAF1* do not result in a phenotype.

The phenotype did not differ significantly between patients with *de novo* and biallelic *DEAF1* variants, except for the presence of microcephaly ($p < 0.0001$). The latter feature was present in three out of the five patients with biallelic variants. Patients with *de novo* or biallelic *DEAF1* variants had DD/ID with severe speech delay, autism, striking mood swings, epilepsy, gait ataxia or other walking problems, a high pain threshold, sleep disturbance, brain abnormalities, and extra-neurologic manifestations, such as recurrent infections and gastrointestinal abnormalities that included feeding difficulties, constipation, and/or diarrhea. This suggests that DAND consist of a phenotypic spectrum of features centered around neurodevelopmental delay rather than two distinct clinical syndromes, but larger number of patients will be required to confirm this hypothesis. To further delineate the clinical spectrum associated with de mutations in *DEAF1*, we established a website to collect detailed clinical information of additional individuals to be identified over the coming years (http://humandiseasegenes.nl/deaf1/). As such, the clinical picture may be quite non-specific. Syndromes to be included in DAND differential diagnosis would be Angelman syndrome, mostly due to ID, gait abnormalities, seizures, paroxysmal bursts of laughter, and fascination with water, and Smith-Magenis syndrome, in

particular due to the combination of ID and severe sleeping problems,[15] and Rett syndrome for the biallelic *DEAF1* variants, for the combination of microcephaly and regression.

The exclusive occurrence of microcephaly and an increased prevalence of some phenotypic abnormalities, such as movement disorder and recurrent structural brain abnormalities, in patients with recessive *DEAF1* variants may also suggest that patients with recessive *DEAF1* variants may present with a different and somewhat more severe clinical condition than patients with *de novo* variants, as these variants have a different effect on a molecular level. One exception to the classical DAND phenotype was observed in one patient with biallelic *DEAF1* variants as he has preserved cognitive abilities. Nevertheless, this patient with variants p.(Asp369Alafs*51) and p.(Cys540Metfs*18) in compound heterozygosity shared clinical features with other patients with biallelic *DEAF1* variants, specifically inability to walk or to speak,[16,18-20] movement disorder,[19] and minor brain abnormalities including corpus callosum, ventricles and white matter abnormalities.[18,20] Another exception to the classical DAND phenotype, is the occurrence of normal speech in one patient with the *de novo* dominant variant p.(Leu214Val), who manifests ID, showing the phenotypic spectrum is wider than previously described.

In conclusion, we propose that different variants in the *DEAF1* gene result in a phenotypic spectrum centered around neurodevelopmental delay. A pathogenic *de novo* dominant variant would incapacitate the product of the wild-type allele and result in dominant-negative effect, whereas a combination of two recessive variants

would result in a partial loss-of-function. Since the clinical picture can be quite non-specific, detailed phenotype information, segregation analysis, and functional analysis are fundamental to determine the pathogenicity of novel variants and to improve the care of these patients.

Acknowledgements

We thank the patients and their parents for participating in this study. We are grateful to Claire Thompson, from the Medical Photography Department, Ninewells Hospital and Medical School, Dundee, Angus (UK) for taking and providing the clinical photographs of individual AD/3. This work was financially supported by grants from the Dutch Organization for Health Research and Development (ZON-MW grants 917-86-319 and 912-12-109 to B.B.A.d.V.), the Cedars-Sinai Diana and Steve Marienhoff Fashion Industries Guild Endowed Fellowship in Pediatric Neuromuscular Diseases (to T.M.P.), the March of Dimes (Grant 6-FY14-422 to M.C.M.), and the National Institutes of Health (Grants NINDS 5R21NS091724 to P.J. and NINDS R01NS069605 to H.C.M.).

References:

1. Vulto-van Silfhout AT, Rajamanickam S, Jensik PJ, et al. Mutations affecting the SAND domain of DEAF1 cause intellectual disability with severe speech impairment and behavioral problems. *Am J Hum Genet.* 2014;94(5):649-661.
2. Yip L, Su L, Sheng D, et al. Deaf1 isoforms control the expression of genes encoding peripheral tissue antigens in the pancreatic lymph nodes during type 1 diabetes. *Nat Immunol.* 2009;10(9):1026-1033.
3. Barker HE, Smyth GK, Wettenhall J, et al. Deaf-1 regulates epithelial cell proliferation and side-branching in the mammary gland. *BMC Dev Biol.* 2008;8:94.
4. Gross CT, McGinnis W. DEAF-1, a novel protein that binds an essential region in a Deformed response element. *EMBO J.* 1996;15(8):1961-1970.
5. Yip L, Creusot RJ, Pager CT, Sarnow P, Fathman CG. Reduced DEAF1 function during type 1 diabetes inhibits translation in lymph node stromal cells by suppressing Eif4g3. *J Mol Cell Biol.* 2013;5(2):99-110.
6. Bottomley MJ, Collard MW, Huggenvik JI, Liu Z, Gibson TJ, Sattler M. The SAND domain structure defines a novel DNA-binding fold in transcriptional regulation. *Nat Struct Biol.* 2001;8(7):626-633.
7. Jensik PJ, Huggenvik JI, Collard MW. Identification of a nuclear export signal and protein interaction domains in deformed epidermal autoregulatory factor-1 (DEAF-1). *J Biol Chem.* 2004;279(31):32692-32699.
8. Jensik PJ, Huggenvik JI, Collard MW. Deformed epidermal autoregulatory factor-1 (DEAF1) interacts with the Ku70 subunit of the DNA-dependent protein kinase complex. *PLoS One.* 2012;7(3):e33404.
9. Kateb F, Perrin H, Tripsianes K, et al. Structural and functional analysis of the DEAF-1 and BS69 MYND domains. *PLoS One.* 2013;8(1):e54715.
10. Hahm K, Sum EY, Fujiwara Y, Lindeman GJ, Visvader JE, Orkin SH. Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1. *Mol Cell Biol.* 2004;24(5):2074-2082.
11. Veraksa A, Kennison J, McGinnis W. DEAF-1 function is essential for the early embryonic development of Drosophila. *Genesis.* 2002;33(2):67-76.
12. Vissers LE, de Ligt J, Gilissen C, et al. A de novo paradigm for mental retardation. *Nat Genet.* 2010;42(12):1109-1112.
13. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet.* 2012;380(9854):1674-1682.
14. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med.* 2017;19(2):209-214.
15. Berger SI, Ciccone C, Simon KL, et al. Exome analysis of Smith-Magenis-like syndrome cohort identifies de novo likely pathogenic variants. *Hum Genet.* 2017;136(4):409-420.
16. Chen L, Jensik PJ, Alaimo JT, et al. Functional analysis of novel DEAF1 variants identified through clinical exome sequencing expands DEAF1-associated

neurodevelopmental disorder (DAND) phenotype. *Hum Mutat.* 2017;38(12):1774-1785.
17. Li SJ, Yu SS, Luo HY, et al. Two de novo variations identified by massively parallel sequencing in 13 Chinese families with children diagnosed with autism spectrum disorder. *Clin Chim Acta.* 2018;479:144-147.
18. Faqeih EA, Al-Owain M, Colak D, et al. Novel homozygous DEAF1 variant suspected in causing white matter disease, intellectual disability, and microcephaly. *Am J Med Genet A.* 2014;164A(6):1565-1570.
19. Rajab A, Schuelke M, Gill E, et al. Recessive DEAF1 mutation associates with autism, intellectual disability, basal ganglia dysfunction and epilepsy. *J Med Genet.* 2015;52(9):607-611.
20. Gund C, Powis Z, Alcaraz W, Desai S, Baranano K. Identification of a syndrome comprising microcephaly and intellectual disability but not white matter disease associated with a homozygous c.676C>T p.R226W DEAF1 mutation. *Am J Med Genet A.* 2016;170A(5):1330-1332.
21. Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. *Nat Genet.* 2004;36(9):949-951.
22. Grimberg J, Nawoschik S, Belluscio L, McKee R, Turck A, Eisenberg A. A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Res.* 1989;17(20):8390.
23. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med.* 2012;367(20):1921-1929.
24. Krumm N, Sudmant PH, Ko A, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res.* 2012;22(8):1525-1532.
25. Pfundt R, Del Rosario M, Vissers L, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet Med.* 2017;19(6):667-675.
26. Jacquemin V, Rieunier G, Jacob S, et al. Underexpression and abnormal localization of ATM products in ataxia telangiectasia patients bearing ATM missense mutations. *Eur J Hum Genet.* 2012;20(3):305-312.
27. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.
28. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812-3814.
29. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;Chapter 7:Unit7 20.
30. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46(3):310-315.
31. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291.
32. Michelson RJ, Collard MW, Ziemba AJ, Persinger J, Bartholomew B, Huggenvik JI. Nuclear DEAF-1-related (NUDR) protein contains a novel DNA binding

domain and represses transcription of the heterogeneous nuclear ribonucleoprotein A2/B1 promoter. *J Biol Chem.* 1999;274(43):30510-30519.

Figure Legends

Figure 1. (a) Structure of the *DEAF1* pre-mRNA [NM_021008.3] with light blue boxes indicating the 5'and 3' untranslated regions (UTRs) and dark blue boxes representing the coding region of exons 1 to 12, connected by lines indicate introns (not drawn to scale). (b) Structure of DEAF1 protein [NP_066288], including the DNA binding and dimerization SAND domain, the zinc-finger homology domain (ZnF) also involved in DNA binding, nuclear localization signal (NLS), nuclear export signal (NES), and MYND domain. Distribution of (c) *de novo* heterozygous variants leading to autosomal dominant mental retardation 24 and (d) homozygous or compound heterozygous variants leading to autosomal recessive dyskinesia, seizures, and intellectual developmental disorder. The number of patients described with the depicted variant are indicated between parenthesis if the variant was indentified in more than one case. Novel variants are indicated in bold. (e) Non-specific facial dysmorphisms of individuals with pathogenic *de novo DEAF1* variants (A - L) and biallelic *DEAF1* variants (M - P), showing common characteristics as horizontal eyebrow, broad nasal tip, exaggerated cupid's bow, thick lower lip vermilion and pointed chin: (A) Individual with p.(Gly212Ser), at 20 years old (AD/1); (B) Individual with p.(Thr213Pro), at 7 years old (AD/3); (C) Individual p.(Leu214Val), at 14 years old (AD/4); (D) Individual with p.(Leu214Pro), at 3 years old (AD/5); (E) Individual with p.(Lys216Glu), at 10 years old (AD/6); (F) Individual with p.(Lys216Asn), at 7 years old (AD/7); (G) Individual with p.(Pro174_Gly222del), at 6 years old (AD/8); (H) Individual with p.(Gly225Glu), at 5 years old (AD/9); (I) Individual with p.(Ser236Gly), at 4 years old (AD/11); (J)

Individual with p.(Arg254del), at 25 years old (AD/13); (K) Individual with p.(Leu272Ser), at 2 years old (AD/15); (L) Individual with p.(Ala276Pro), at 14 years old (AD/17); (M) Individual with p.(Asp369Alafs*51) and p.(Cys540Metfs*18), at 17 years old (AR/1); (N) Individual with p.(Trp234*) and p.(Glu239Gly), at 16 years old (AR/3); (O) Individual with p.(Arg224Gln) in homozygosity, at 10 years old (AR/4); (P) Individual with p.(Arg224Gln) in homozygosity), at 2 years old (AR/5). Individuals with *de novo* variants may have high forehead and/or macrocephaly, while individuals with biallelic variants may have narrow forehead and microcephaly.

Figure 2. The *de novo* variant c.664+1G>T results in an in-frame deletion within the *DEAF1* SAND domain. (A) Location of the c.664+1G>T variant relative to exons 2-5 in the *DEAF1* gene. (B) PCR using primers that would amplify exons 2-5 of *DEAF1* mRNA (gray bars shown in A) with cDNA from control or c.664+1G>T EBV-LCL (individual AD/8) treated with vehicle (-) or CHX (+). A single DNA band was observed in control EBV-LCL corresponding to correctly spliced DEAF1. An additional, approximately 150 basepairs smaller, PCR product was identified in c.664+1G>T EBV-LCL. (C) DNA sequencing indicated c.664+1G>T EBV-LCL have both correctly spliced DEAF1 RNA (above) as well as DEAF1 mRNA that lack Exon4 (below) which could result in a truncated p.Pro174_Gly222del DEAF1 protein. (D) Lysates from HEK293t cells transfected with the expression constructs for WT (~80 kDa) or p.(Pro174_Gly222del) variant (~75 kDa) were separated by SDS-PAGE and analyzed by Western blot using DEAF1 (red) and b-ACTIN (green) antibodies. (E)

Lysates from control or c.664+1G>T EBV-LCL were immunoprecipitated with preimmune serum (PI) or DEAF1 antibodies followed by Western blots with DEAF1 antibodies. The control EBV-LCL shows a single 80 kDa full length DEAF1 protein band. An additional ~75 kDa DEAF1 protein was also observed in the EBV-LCL with the c.664+1G>T variant. Closed arrows indicate DEAF1 protein bands and open arrow indicates heavy chain IgG (IgG HC).

Figure 3. Effect of *DEAF1* variants on transcriptional activity. (A and B) *DEAF1* promoter activity after transfection of HEK293t cells with WT or the indicated variants. (A) Compared to WT, all *de novo* variants resulted in loss of transcriptional repression activity at the *DEAF1* promoter. (B) Heritable biallelic variants had no effect on *DEAF1* promoter transcriptional repression, except for p.(Trp234*). (C and D) *Eif4g3* promoter activity after transfection of HEK293t cells with WT or the indicated variants. (C) All *de novo* variants, except p.(Pro293Leu) and p.(Ser236Gly), resulted in suppression of transcription relative to WT transactivation. No change from basal promoter activity (pcDNA3) was detected for p.(Ser236Gly) and increased promoter activity was observed for p.(Pro293Leu). (D) Heritable biallelic variants had no effect on *Eif4g3* promoter activation compared to WT. Each bar represents the mean +/- SEM of the normalized luciferase activity of three independent experiments when the activity of pcDNA3 (DEAF1 promoter alone) was set to 100% (N= 5-7 *de novo* variants with N=14-15 WT, N=4-5 heritable variants with N=10-12 WT). **p<0.01 One-way ANOVA with Dunnett's multiple comparison WT DEAF1 versus each mutant.