Genetics in Medicine (2022) 24, 2051–2064

ELSEVIER

Genetics in Medicine

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Missense variants in *ANKRD11* cause KBG syndrome by impairment of stability or transcriptional activity of the encoded protein

ARTICLE INFO

Article history:
Received 27 January 2022
Received in revised form
20 June 2022
Accepted 21 June 2022
Available online 14 July 2022

Keywords:
ANKRD11
Genotype–phenotype study
KBG syndrome
Missense variants
Neurodevelopmental disorders

ABSTRACT

Purpose: Although haploinsufficiency of *ANKRD11* is among the most common genetic causes of neurodevelopmental disorders, the role of rare *ANKRD11* missense variation remains unclear. We characterized clinical, molecular, and functional spectra of *ANKRD11* missense variants.
Methods: We collected clinical information of individuals with *ANKRD11* missense variants and evaluated phenotypic fit to KBG syndrome. We assessed pathogenicity of variants through *in silico* analyses and cell-based experiments.
Results: We identified 20 unique, mostly de novo, *ANKRD11* missense variants in 29 individuals, presenting with syndromic neurodevelopmental disorders similar to KBG syndrome caused by *ANKRD11* protein truncating variants or 16q24.3 microdeletions. Missense variants significantly clustered in repression domain 2 at the ANKRD11 C-terminus. Of the 10 functionally studied missense variants, 6 reduced ANKRD11 stability. One variant caused decreased proteasome degradation and loss of ANKRD11 transcriptional activity.
Conclusion: Our study indicates that pathogenic heterozygous *ANKRD11* missense variants cause the clinically recognizable KBG syndrome. Disrupted transrepression capacity and reduced protein stability each independently lead to ANKRD11 loss-of-function, consistent with haploinsufficiency. This highlights the diagnostic relevance of *ANKRD11* missense variants, but also poses diagnostic challenges because the KBG-associated phenotype may be mild and inherited pathogenic *ANKRD11* (missense) variants are increasingly observed, warranting stringent variant classification and careful phenotyping.

Introduction

KBG syndrome (OMIM 148050) is an autosomal dominant neurodevelopmental disorder (NDD) typically characterized by mild intellectual disability (ID) or developmental delay, macrodontia of upper central permanent incisors, mild skeletal anomalies, behavioral disturbances, and distinctive craniofacial features.[1-6] Although KBG syndrome is considered a clinically recognizable syndrome with macrodontia as its most defining trait,[6] there is considerable clinical variability and none of the KBG features is pathognomonic. Hence, despite being described as a clinical entity since 1975,[1] KBG syndrome was underdiagnosed before causative *ANKRD11* variants were discovered.[7] The exact prevalence of KBG syndrome is not established but it is thought to be a relatively common cause of genetic NDDs, with the

Elke de Boer and Charlotte W. Ockeloen contributed equally.
Maggie M.K. Wong and Tjitske Kleefstra contributed equally.
*Correspondence and requests for materials should be addressed to Charlotte W. Ockeloen, Department of Genetics, Radboudumc, P.O. Box 9101, 6500 HB, Nijmegen The Netherlands. *E-mail address:* Charlotte.Ockeloen@radboudumc.nl
A full list of authors and affiliations appears at the end of the paper.

doi: https://doi.org/10.1016/j.gim.2022.06.007

associated gene (*ANKRD11*) in the top 3 of mutated genes in NDD cohorts accounting for 0.5% to 1% of diagnoses.[8,9]

KBG syndrome is caused by heterozygous protein truncating variants (PTVs) in *ANKRD11* (encoding ANKRD11) or by 16q24.3 microdeletions encompassing (part of) *ANKRD11*. PTVs and microdeletions explain all cases in 4 previously described KBG cohorts,[2-5] whereas in the general population, *ANKRD11* shows strong constraint against loss-of-function variation (probability of being loss-of-function intolerant [pLI] = 1; observed/expected [o/e] = 0.05 [0.02-0.11]; gnomAD v2.1.1).[10] Therefore, haploinsufficiency of *ANKRD11* is commonly accepted as mechanism of pathogenicity for KBG syndrome.[11] This is supported by observations of reduced amounts of *ANKRD11* messenger RNA and protein when the gene contains a PTV,[12] suggesting that variants trigger the nonsense-mediated decay (NMD) pathway,[12] although PTVs leading to (partial) escape from NMD have also been described.[7,13] Also consistent with haploinsufficiency is the finding that ANKRD11 mutated with p.(Lys1347del) or p.(Leu2143Val) shows reduced transcriptional activity on the *p21* promotor in cell-based systems, that can be rescued by wild type but not mutated ANKRD11.[14]

ANKRD11 is ubiquitously expressed and localizes mainly to the nucleus in a homogenous pattern. ANKRD11 is a crucial regulator of neuronal development[7,15] that interacts with coactivators and corepressors of transcription,[16] showing (co) regulatory effects on various sets of genes. These include genes encoding signaling molecules, chromatin remodelers, and transcriptional regulators,[15] controlling histone acetylation and gene expression during neural development. ANKRD11 contains 3 transcriptional regulatory domains: 1 activation domain and 2 repression domains (RDs). The RDs, located at the N-terminus (RD1) and C-terminus of ANKRD11 (RD2), functionally outweigh the activation domain, because full-length ANKRD11 functions as a repressor of ligand-dependent transcription.[17] Interaction of ANKRD11 with other proteins and homodimerization are mediated through ankyrin repeats, located at the N-terminus.[13] The C-terminal part of ANKRD11, containing (predicted) destruction box motifs (D-boxes), was suggested to be critical for its degradation.[13]

Whereas PTVs in *ANKRD11* are a well-recognized cause of KBG syndrome, the role of rare missense variants remains ambiguous. Contrary to what is seen for PTVs, constraint metrics based on the general population indicate that missense variants tend to be well-tolerated (z-score –0.55; o/e = 1.04 [1-1.08]).[10] There are numerous entries of *ANKRD11* missense variants in ClinVar (access date August 20, 2021) but only approximately 6% are classified as (likely) pathogenic and almost half as variants of uncertain significance.[18] In the literature, approximately 2.6% of (de novo) variants in *ANKRD11* are missense variants,[19] listed in Supplemental Table 1. Missense variants are reported with varying levels of evidence on pathogenicity, and functional studies have only been performed for p.(Leu2143Val), showing a loss-of-function effect.[14] The Yoda mutant mouse (C3H.Cg-Ankrd11Yod/H, p.[Glu2502Lys]) carries an *Ankrd11* missense variant and shows phenotypic overlap with core features of KBG syndrome, including reduced body size and craniofacial abnormalities such as shortened snouts with deformed nasal bones, wider skulls, and failure of cranial sutures to close.[20] In addition, Yoda mice show behavioral abnormalities reflective of cognitive dysfunction.[15] On the cellular level, the heterozygous Yoda variant causes similar cellular perturbations of abnormal neuronal precursor proliferation and localization of neurons as seen for Ankrd11 knockdown,[15] suggesting a loss-of-function mechanism. However, a dominant-negative mechanism has also been hypothesized to contribute to the Yoda mouse phenotype.[13] Ankrd11 was shown to mislocalize to the nucleolus, possibly resulting from diminished degradation.[13] Because the N-terminal ankyrin repeats are unaffected by the variant, dimerization of wild-type and mutant Ankrd11 was hypothesized to result in decreased degradation of both proteins, potentially implicating such dominant-negative mechanism.[13] So far, a dominant-negative mechanism has not been confirmed in additional studies. In general, consequences of *ANKRD11* missense variants on clinical phenotypes and protein function are largely unknown.

We characterized genotypes, phenotypes, and functional consequences associated with *ANKRD11* missense variants by describing a cohort of 29 individuals. Most individuals exhibit both characteristic facial appearance and other KBG-associated features, fitting well within the clinical spectrum described for KBG syndrome. We showed that missense variants in *ANKRD11* significantly cluster in the C-terminal RD2, with an overrepresentation of mutated arginine residues. Missense variants result in a loss of normal ANKRD11 function, either caused by reduced protein stability with normal or increased proteasome degradation or caused by a loss of transrepression capacity with decreased proteasome degradation. Our findings are consistent with *ANKRD11* haploinsufficiency, mechanistically underlying KBG syndrome caused by PTVs or 16q24.3 microdeletions.

Materials and Methods

Clinical and *In Silico* characterization

Identification and clinical characterization of individuals and *in silico* analyses of (likely) pathogenic *ANKRD11* missense variants are described in the Supplementary Methods. Variants were annotated in the context of genome build GRCh37/Hg19, using transcript reference sequence NM_013275.6 and protein reference sequence NP_037407.4. Human Phenotype Ontology (HPO)-based[21] clustering was performed as previously described[22] using clinical data of 29 individuals with *ANKRD11* missense variants and 35 individuals with *ANKRD11* PTVs or microdeletions (Supplemental Table 2A-C, Supplemental JSON) after grouping HPO-data on the basis of semantic similarity (Supplemental Table 3A).[23,24] Spatial clustering of independently observed missense variants (25/29) was performed as previously described,[25] excluding 4 familial variants. The Supplementary Methods contain details of both clustering analyses. P values < .05 were considered significant.

Missense permutation analysis

To test whether the observed number of variants affecting arginine residues was significantly greater than expected by chance, *ANKRD11* (ENST00000301030.10/NM_013275.6) was mutated *in silico* and output was annotated using Ensembl Variant Effect Predictor v104.[26] To generate an expected missense distribution, sets of 17 missense variants in RD2 and 8 missense variants outside RD2 (based on 25 independently observed missense variants, excluding 4 familial variants) were randomly sampled 100,000 times using per-nucleotide mutation rates as weights.[27] Number of missense variants affecting arginine residues inside and outside RD2 were counted per iteration. *P* values were computed using a permutation test by ranking the observed number of variants affecting arginine residues within the set of 100,000 expected values. *P* values < .05 were considered significant.

Cell culture and transfection

HEK293T/17 cells (CRL-11268, ATCC) were cultured in Dulbecco's Modified Eagle Medium (Gibco) supplemented with 10% fetal bovine serum (Gibco) and 100 U/mL penicillin-streptomycin (Thermo Fisher Scientific) at 37°C and 5% carbon dioxide. For immunofluorescence analysis, cells were seeded onto coverslips coated with 100 µg/mL poly-D-lysine (Merck, Millipore). Transfections were performed using GeneJuice (Merck, Millipore) following the manufacturer's instructions or polyethyenimine in 3:1 ratio with total mass of DNA transfected.

DNA constructs and site-directed mutagenesis

Full-length wild-type ANKRD11 construct fused to a C-terminal Myc-DDK tag under a human cytomegalovirus (CMV) promoter (pCMV-Entry-ANKRD11) was purchased from Origene (RC211717). To generate an N-terminal EGFP-tag, sequence encoding EGFP was cut from a pEGFP-C2 vector (Clontech) and subcloned into the pCMV-Entry-ANKRD11 plasmid using KpnI/NdeI restriction sites. Constructs (pCMV-Entry-EGFP-ANKRD11) carrying (in frame) *ANKRD11* missense variants were generated using a QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent) following the manufacturer's protocol. All constructs were verified using Sanger sequencing. Supplemental Table 4 lists primer sequences.

Fluorescence imaging of subcellular localization

HEK 293T/17 cells grown on poly-D-lysine-coated coverslips were transiently transfected with 500 ng of pCMV-EGFP-ANKRD11 constructs 24 hours after seeding. Cells were fixated 48 hours after transfection using 4% paraformaldehyde solution (Electron Microscopy Supplies Ltd) for 20 minutes at room temperature. Hoechst 33342 (Invitrogen) was used for nuclear staining, before mounting with VECTASHIELD Antifade Mounting Medium (Vectorlab). Fluorescence images were obtained using an LSM880 AxioObserved confocal microscope (Zeiss). For images of single nuclei, the Airyscan unit (Zeiss) was used with a 4.0 zoom factor. Images were analyzed using the ImageJ "Analyze particle" plugin.

Fluorescence-based quantification of protein stability and degradation

HEK293/T17 cells were transfected in triplicate in clear-bottomed black 96-well plates with EGFP-tagged ANKRD11 variants. After 48 hours, cycloheximide (Sigma Aldrich) at 50 µg/mL or MG132 (R&D Systems) at 5 µg/mL was added. Cells were incubated at 37°C with 5% carbon dioxide in the Infinite M200PRO microplate reader (Tecan), and fluorescence intensity of EGFP (excitation: 503 nm, emission: 540 nm) was measured over 24 hours at 3-hour intervals. Statistical analysis was performed using 2-way analysis of variance followed by Dunnett's correction for multiple testing. *P* values < .05 were considered significant.

Luciferase reporter assays

We used firefly luciferase reporters: pGL2-p21 promoter-Luc and WWP-Luc carrying the promoter region of *CDKN1A/P21*. Reporters were gifts from Martin Walsh (Addgene plasmid #33021; http://n2t.net/addgene:33021; RRID:Addgene_33021)[28] and Bert Vogelstein (Addgene plasmid 16451; http://n2t.net/addgene:16451; RRID:Addgene_16451),[29] respectively. HEK293/T17 cells were seeded in clear-bottomed white 96-well plates (Greiner Bio-One). Cells were cotransfected with 320 ng of firefly luciferase reporter construct, 6.5 ng of pGL4.74 Renilla luciferase normalization control, and 1000 ng of an EGFP-ANKRD11 expression construct or empty EGFP expression vector. After 48 hours, firefly luciferase and Renilla luciferase activities were measured using a Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's instructions on an Infinite F Plex microplate reader (Tecan). Statistical analysis was performed using a 1-way analysis of variance followed by Dunnett's post hoc test. *P* values < .05 were considered significant.

Results

ANKRD11 missense variants cause syndromic neurodevelopmental phenotypes

Through international collaborations[30,31] we identified 29 individuals with rare missense variants in *ANKRD11* (Figure 1; Supplemental Table 2A). The cohort consisted

of 24 unrelated individuals with *ANKRD11* missense variants—of which 20 variants occurred de novo—and 1 family with 5 affected individuals from 3 generations (Supplemental Figure 1). For 5 individuals, inheritance status of the variant could not be established. The cohort comprises 18 males and 11 females, with an age range of 7 months to 73 years. We observed syndromic neurodevelopmental phenotypes, summarized in Table 1 with detailed data compiled in Supplemental Table 2A and B and Supplementary Results. Most frequent phenotypic features were facial dysmorphisms (28/29 [96.6%]; fitting characteristic dysmorphisms of KBG, 23/29 [79.3%]; Figure 2A), behavioral disturbances (25/28 [89.3%]), neurodevelopmental delay (26/28 [92.9%]; speech delay, 23/26 [88.5%]; motor delay, 20/27 [74.1%]), mild to moderate ID (22/27 [81.5%]; borderline, 2/27 [7.4%]; mild, 15/27 [55.6%]; moderate, 4/27 [14.8%]; unknown severity, 3/27 [11.1%]), abnormal dentition (21/26, [80.8%]; macrodontia of upper central incisors, 13/24 [54.2%]; other dental abnormalities, 14/21 [66.7%]; Figure 2B), and hand abnormalities (20/25 [80%]; Figure 2C). Of note, individual 19 also carried a likely pathogenic variant in *ARID2*, implicated in Coffin-Siris syndrome 6 (OMIM 617808), and in the family of 5 individuals an additional *ANKRD11* variant of uncertain significance, p.(Pro61Ser), was observed in cis with the pathogenic p.(Arg2579His) variant.

ANKRD11 missense variants are predicted deleterious and cluster at C-terminal RD2

We found 20 unique missense variants (Figure 1; Supplemental Table 5) significantly clustering at the highly intolerant C-terminal RD2 ($P = 9.99e-9$), with recurrence of p.(Arg2512Gln), p.(Glu2522Lys), p.(Arg2579His), p.(Arg2585Cys), and p.(Leu2605Arg). In addition, the arginine residues at p.2512, p.2536, and p.2579 were affected by 2 different missense variants, and p.(Glu2522Lys) is equivalent to the orthologous *Ankrd11* p.(Glu2502Lys) in the Yoda mouse (Supplemental Figure 2). Most of observed variants are predicted deleterious and affect conserved and intolerant residues (Figure 1B and C), with no predicted effects on premessenger RNA splicing (Supplemental Table 5). Two variants, p.(Glu2522Lys) and p.(Arg2523Trp), affect a ProViz[32] predicted (low consensus similarity) D-box (Supplemental Table 6; Supplemental Figure 3) and 2 variants, p.(Arg2512Leu) and p.(Arg2512Gln), are located at an additional RxxL-motif, marked as D-box in literature.[13] We observed a significantly greater number of missense variants affecting arginine residues within RD2 (12/17 variants [70.6%]) than would be expected by chance ($P = 1.00e-4$), whereas such enrichment was not observed outside RD2 (0/8 variants; 0%; $P = .37$), visualized in Supplemental Figure 4. We therefore hypothesized that generally, *ANKRD11* missense variants affecting arginine residues in RD2 are likely to be pathogenic.

Phenotypes associated with *ANKRD11* missense variants fit the KBG-associated clinical spectrum

On the basis of 4 large published cohorts together describing 135 individuals with KBG syndrome caused by PTVs or 16q24.3 microdeletions affecting *ANKRD11*,[2-5] we assessed phenotypic fit of each of the individuals in the missense cohort to the clinical spectrum associated with KBG syndrome. The majority of individuals (23/29 [79.3%]) exhibited dysmorphisms fitting the characteristic facial gestalt by which KBG syndrome can be recognized (Figure 2A and B; Supplemental Table 2A and B, Supplemental Table 5). Indeed, 2 individuals in the cohort (individual 6 and 14) were diagnostically evaluated through targeted *ANKRD11* Sanger sequencing because of a high clinical suspicion. Of the 29 individuals, 22 (75.9%) met the diagnostic criteria described for KBG syndrome[4] (Supplemental Table 2A). After review of the observed phenotypes by expert clinicians, it was concluded that almost all individuals fit the KBG-associated phenotypic spectrum. These included many individuals not fully meeting the diagnostic criteria, either because some features of KBG syndrome not captured in the diagnostic criteria (eg, delayed bone age and congenital heart defects) were seen or because the characteristic facial appearance of KBG syndrome was observed. Only for individuals 3 and 29, carrying p.(Leu509Pro) and p.(Leu2605Arg), phenotypic fit to the clinical spectrum of KBG syndrome was considered poor.

We next compared the group of individuals with missense variants with the collective KBG cohorts[2-5] and found that frequencies of most KBG-associated features observed in the missense cohort lie within the range of frequencies of these features seen in the group of individuals with PTVs or microdeletions (Table 2). We therefore hypothesized that KBG syndrome resulting from missense variants is indistinguishable from KBG syndrome caused by *ANKRD11* PTVs or microdeletions.

To quantitatively investigate this hypothesis, we compared standardized clinical data (Supplemental JSON[21]) of 29 individuals with missense variants and 35 individuals with KBG syndrome caused by *ANKRD11* PTVs or microdeletions affecting *ANKRD11* only. Applying a Partitioning Around Medoids clustering algorithm[33] on 68 features derived from HPO-data resulted in correct classification of 40 of 64 individuals as either belonging to the PTV or missense variant group ($P = .04396$; Figure 2D; Supplemental Table 3A-C), indicating that the algorithm recognizes a difference between the group with PTVs or microdeletions and the group with missense variants. This challenges the clinical observation that pathogenic *ANKRD11* missense variants and PTVs or microdeletions cause the same clinical entity.

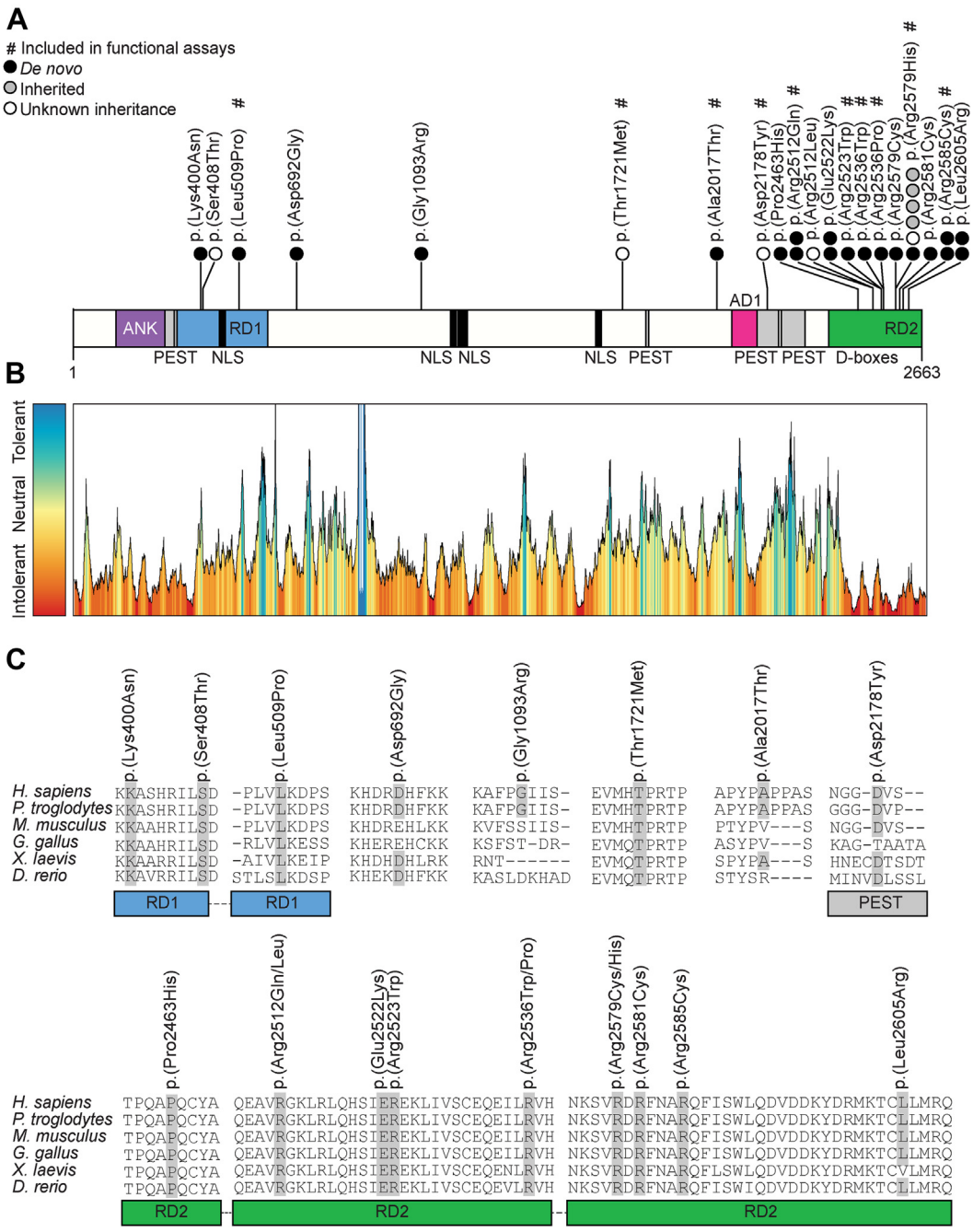

Figure 1 Missense variants cluster in the intrinsic repressor domain 2 in the C-terminus of the ANKRD11 protein. A. Schematic representation of ANKRD11 (UniProt: Q6UB99, having 100% sequence identity with NP_037407.4) indicating the location of variants included in this study. De novo variants are indicated by black disks, inherited variants by gray disks, variants with unknown inheritance by white disks, and variants marked with # are included in functional assays. The ANKRD11 protein sequence (2663 amino acids) contains an ankyrin-repeat domain (ANK; purple; amino acids 133-296), 4 PEST sequences (gray; amino acids 286-324, 1796-1806, 2146-2212, and 2222-2297), 4 bipartite nuclear localization signals (NLS; black; amino acids 1184-1200, 1208-1236, 1358-1374, 1640-1656), 2 intrinsic repressor domains (RD) (RD1; blue; amino acids 318-611 and RD2; green; amino acids 2369-2663), and an activator domain (AD1; pink; amino acids 2076-2145). An overview with variant details per subject is provided in Supplemental Table 2A, with details on variant interpretation in Supplemental Table 5. B. MetaDome analysis of the ANKRD11 missense variants. Overview of the ANKRD11 protein (NP_037407.4) tolerance landscape visualized via the MetaDome web server version 1.0.1. The green and blue peaks correspond to regions more tolerant to missense variation and the red valleys indicate intolerant regions. C. Sequence alignment of the region containing part of ANKRD11 amino acid sequence in human (UniProt: Q6UB99), chimpanzee (A0A2I3TR65), mouse (E9Q4F7), chicken (A0A3Q2UE98), African clawed frog (A0A1L8GEN1), and zebrafish (E7F5R3). Residues in which missense variants were found are highlighted in gray. D-boxes, destruction box motifs.

Table 1 Summary of observed clinical features

Features Associated With KBG Syndrome	%	Present/Total Assessed
Macrodontia upper central incisors[2]	54.2	13/24
Additional dental abnormalities[2]	66.7	14/21
Fitting the KBG characteristic facial appearance[2]	79.3	23/29
Hand abnormalities[2]	80	20/25
Postnatal short stature[2] (<−2 SD)	53.6	15/28
Delayed bone age[2]	57.1	8/14
Costovertebral anomalies[2]	7.1	1/14
ID[2]	81.5	22/27
Normal IQ	11.1	3/27
ID of unknown severity	11.1	3/27
Borderline	7.4	2/27
Mild	55.6	15/27
Moderate	14.8	4/27
Developmental delay[2]	92.9	26/28
Seizures[2]	22.2	6/27
Behavioral abnormalities[2]	89.3	25/28
ADHD/hyperactive behavior	69.2	18/26
Autism spectrum disorder	36	9/25
Anxiety	37.5	9/24
Cryptorchidism[2]	20	3/15
Congenital heart defect[2]	32	8/25
Palate defect[2]	7.1	2/28
Hearing loss[2]	39.3	11/28
First degree relative with KBG syndrome[2]	19.2	5/26
Large fontanelle at birth[4]	18.2	2/11
Feeding difficulties[4]	33.3	9/27
Precocious puberty[3]	5.9	1/17
Meeting diagnostic criteria of KBG syndrome[4]	75.9	22/29
Other features		
Motor delay	74.1	20/27
Speech delay	88.5	23/26
Hypotonia	41.7	10/24
Hypertonia/spasticity	12	3/25
Sleep disturbances	30.8	8/26
Abnormal brain MRI	20	3/15
Abnormalities during pregnancy	45.8	11/24
Abnormalities during delivery	57.7	15/26
Macrocephaly (>2 SD)	3.7	1/27
Microcephaly (<−2 SD)	22.2	6/27
Vision abnormalities	50	13/26
Gastrointestinal abnormalities	32	8/25
Endocrine/metabolic abnormalities	25.9	7/27
Immunologic abnormalities	18.5	5/27
Skin/hair/nail abnormalities	57.7	15/26

ADHD, attention deficit hyperactivity disorder; *ID*, intellectual disability; *IQ*, intelligence quoteint; *MRI*, magnetic resonance imaging.

ANKRD11 missense variants act via 2 distinct loss-of-function mechanisms

We continued by studying the functional consequences on protein localization, protein stability, proteasome degradation, and transcriptional activity of a subset of the observed *ANKRD11* missense variants using HEK293T/17 cells transiently transfected with mutant ANKRD11. To obtain a comprehensive insight of the spectrum of variants, we examined variants spread across *ANKRD11*, including p.(Leu509Pro) located in RD1; p.(Thr1721Met), p.(Ala2017Thr), and p.(Asp2178Tyr) outside known protein domains; and several variants (p.(Arg2512Gln), p.(Arg2523Trp), p.(Arg2536Trp), p.(Arg2536Pro), p.(Arg2579His), and p.(Arg2585Cys)) in RD2.

We first assessed the effect of missense variants on subcellular localization of ANKRD11. When transiently expressed as EGFP-fusion proteins in HEK239T/17 cells, wild-type ANKRD11 localized to the nucleus in a homogeneous speckle-like pattern, consistent with previous findings in non-neuronal cell lines[16] and mouse neocortical neurons overexpressing wild-type Ankrd11.[7] None of the tested missense variants affected nuclear localization of ANKRD11 (Figure 3A, Supplemental Figure 5), contrasting with the nucleolar mislocalization of mutant Ankrd11 reported in Yoda mice.[13]

On the basis of spatial clustering at the C-terminus, which is critical for ANKRD11 degradation,[13] we hypothesized that missense variants might alter ANKRD11 stability, possibly via altered proteasome degradation. Moreover, 4 variants are located at putative destruction motifs (Supplemental Table 6, Supplemental Figure 3). To assess ANKRD11 protein stability, we treated HEK239T/17 cells expressing EGFP-tagged ANKRD11 with cycloheximide to inhibit translation and measured relative fluorescence intensity over 24 hours. We found that all variants in RD2 except p.(Arg2585Cys) showed reduced protein stability compared with wild type, whereas among variants outside the RD2, only p.(Ala2017Thr) was less stable (Figure 3B, Supplemental Figure 6 and 7). All other variants outside RD2 showed stability similar to wild type. We next examined the effect of missense variants on proteasome-mediated degradation after treating EGFP-ANKRD11 expressing cells with proteasome inhibitor MG132. Only 2 variants, both located in RD2, affected proteasome degradation, showing opposite directions of effect; whereas p.(Arg2523Trp) displayed increased proteasome degradation, p.(Arg2585Cys) showed decreased proteasome degradation (Figure 3B).

To study the effects of missense variation on transcriptional activity of ANKRD11, we performed luciferase reporter assays with the *CDKN1A/P21* promoter, a known downstream target. Of all tested variants, only p.(Arg2585Cys) affected transcriptional activity, leading to a loss of transcriptional repression on *CDKN1A/P21* (Figure 3C and D).

Taken together, our cell-based assays indicate that most missense variants yield a loss-of-function effect, either resulting from a reduced dosage of ANKRD11 due to decreased protein stability with or without increased proteasome degradation or through a loss of transrepressive activity. Three variants, p.(Leu509Pro), p.(Thr1721Met), and p.(Asp2178Tyr), did not result in aberrations in any of the tested protein functions and are therefore classified as of

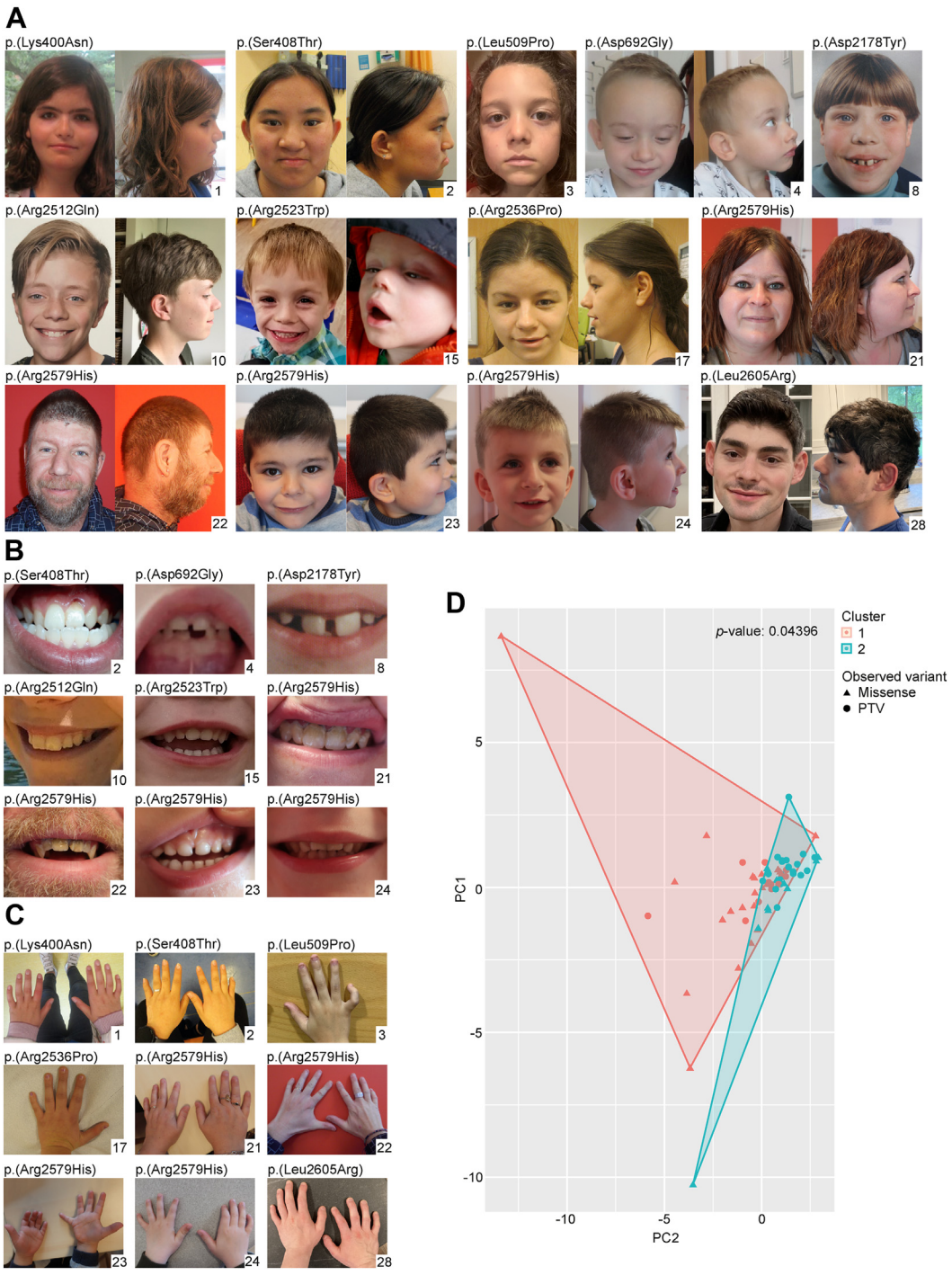

Figure 2 **Clinical evaluation of individuals with *ANKRD11* missense variants.** A. Facial photographs of affected individuals. Features of most individuals fit the characteristic facial gestalt of KBG syndrome (individuals 1, 2, 8, 10, 15, 17, 23, 24, 28). Individuals display a triangular face (individuals 4, 10, 15, 23, 24, 28), with full eyebrows and synophrys (individuals 1, 23, 28). Both downslant (individuals 3, 10) and upslant (individuals 1, 2, 8, 15, 21, 23, 28) of palpebral fissures can be observed. Individuals 2, 4, 8, 10, 15, 21, and 23 exhibit a low nasal bridge, with an upturned nasal tip and triangular-shaped nasal tip in individuals 4, 8, 10, 15, 17, 23, and 24. In individuals 1, 2, 4, 10, 15, and 23 a smooth philtrum is observed, with a thin upper lip in individuals 1, 2, 4, 15, 23, and 24. From the side, low-set, simple ears (individuals 2, 4, 23, 24, 28) and a flat facial profile (individuals 1, 2, 23, 24) can be observed. B. Photographs of teeth, showing macrodontia of central permanent incisors in individuals 8, 10, and 21, with teeth discolorization in individuals 21 and 22, and widely spaced teeth in individuals 4, 8, and 23. C. Photographs of hands, showing short palms (individuals 1, 2, 3, 17, 21, 22, 23, 28) and short fingers (individuals 21, 22, 24, 28) with tapering (individual 2, 21, 22, 23, 24) and clinodactyly of the fifth digit (individual 1, 2, 3, 22, 24, 28). D. Visualization of Human Phenotype Ontology-based clustering analysis, indicating that the clustering algorithm recognizes a difference between individuals with missense variants and individuals with PTVs or 16q24.3 microdeletions. PC, principal component; PTV, protein truncating variant.

Table 2 Comparison of clinical features between individuals with *ANKRD11* missense variants and previously reported KBG cohorts

Publication	Ockeloen et al[2]		Low et al[4]		Goldenberg et al[3]		Gnazzo et al[5]		Cumulative Frequencies in Published Cohorts[2-5]			This Publication	
Variant Type	PTVs + Microdeletions		PTVs		PTVs + Microdeletions		PTVs + Microdeletions		PTVs + Microdeletions			Missense Variants	
Features Associated With KBG Syndrome	%	Present/Total Assessed	%	Present/Total Assessed	%	Present/Total Assessed	%	Present/Total Assessed	%	Present/Total Assessed	Frequency Range (%)	%	Present/Total Assessed
Macrodontia upper central incisors[2]	75.8	25/33	85.2	23/27	69.2	18/26	76.7	23/30	76.7	89/116	69.2-85.2	54.2	13/24
Additional dental abnormalities[2]	48.5	16/33	NA	NA	NA	NA	NA	NA	48.5	16/33	48.5	66.7	14/21
Fitting characteristic facial appearance of KBG[2]	93.9	31/33	40.6	13/32	100	39/39	100	31/31	84.4	114/135	40.6-100	79.3	23/29
Hand abnormalities[2]	81.8	27/33	46.9	15/32	69.7	23/33	67.7	21/31	66.7	86/129	46.9-81.1	80	20/25
Postnatal short stature[2]	54.5	18/33	28.1	9/32	40.5	15/37	58.1	18/31	45.1	60/133	28.1-58.1	53.6	15/28
Delayed bone age[2]	15.2	5/33	60	3/5	66.7	8/12	NA	NA	32	16/50	15.2-66.7	57.1	8/14
Costovertebral anomalies[2]	30.3	10/33	NA	NA	NA	NA	NA	NA	30.3	10/33	30.3	7.1	1/14
ID[2]	60.6	20/33	NA	NA	60	9/15	59.1	13/22	60	42/70	59.1-60.6	81.5	22/27
Learning disability or developmental delay[2]	93.9	31/33	100	32/32	94.3	33/35	96.8	30/31	96.2	126/131	93.9-100	92.9	26/28
Seizures[2]	27.3	9/33	43.8	14/32	31.6	12/38	16.1	5/31	29.9	40/134	16.1-43.8	22.2	6/27
Behavioral abnormalities[2]	78.8	26/33	100	30/30	51.4	19/37	NA	NA	75	75/100	51.4-100	89.3	25/28
Cryptorchidism[2]	40	8/20	31.3	5/16	18.8	3/16	13.3	2/15	26.9	18/67	13.3-40	20	3/15
Congenital heart defect[2]	18.2	6/33	12.5	4/32	25.6	10/39	35.5	11/31	23	31/135	12.5-35.5	32	8/25
Palate defect[2]	21.2	7/33	12.5	4/32	5.1	2/39	0	0/31	9.6	13/135	0-21.2	7.1	2/28
Hearing loss[2]	24.2	8/33	25	8/32	30.6	11/36	19.4	6/31	25	33/132	19.4-30.6	39.3	11/28
First degree relative with KBG syndrome[2]	52.6	10/19	21.9	7/32	20.5	8/39	6.5	2/31	22.3	27/121	6.5-52.6	19.2	5/26
Large fontanelle at birth[4]	NA	NA	21.9	7/32	NA	NA	NA	NA	21.9	7/32	21.9	18.2	2/11
Feeding difficulties[4]	NA	NA	31.3	10/32	NA	NA	35.5	11/31	33.3	21/63	31.3-35.3	33.3	9/27
Precocious puberty[3]	NA	NA	NA	NA	15.6	5/32	5.6	1/18	12	6/50	5.6-15.6	5.9	1/17

ID, intellectual disability; *NA*, not available; *PTV*, protein truncating variant.

uncertain significance, although the individuals carrying p.(Thr1721Met) and p.(Asp2178Tyr) exhibit typical KBG features (individual 6 and 8, respectively). All evidence per observed missense variant including classification on the basis of American College of Medical Genetics and Genomics/Association for Molecular Pathology criteria are provided in Supplemental Table 5.

Discussion

Although KBG syndrome has been clinically recognized for almost 50 years[1] and PTVs and microdeletions affecting *ANKRD11* have been robustly implicated in its etiology since 2011,[7] the role of rare missense variation in *ANKRD11* remained unclear. We characterized clinical, molecular, and functional spectra of *ANKRD11* missense variants by collecting information for 29 individuals and assessing effects of missense variation on ANKRD11 functions. We show that almost all individuals carrying rare heterozygous predicted damaging *ANKRD11* missense variants fit well within the clinical spectrum described for KBG syndrome. Missense variants mainly affect the C-terminal RD2 with an overrepresentation of mutated arginine residues. Based on cellular assays, missense variants result in loss-of-function of ANKRD11, either by impaired protein stability or reduced transcriptional activity, consistent with *ANKRD11* haploinsufficiency causing KBG syndrome through PTVs and microdeletions.

Most individuals presented with characteristics fitting the KBG-associated phenotypic spectrum and, from a clinical perspective, individuals with KBG syndrome caused by *ANKRD11* missense variants or by PTVs or microdeletions are indistinguishable. However, unexpectedly, HPO-based clustering analysis showed a difference between the groups. Possibly, ascertainment bias influenced this analysis, because recognizing pathogenicity for missense variants is more challenging than that for PTVs. In addition, of the 7 individuals with missense variants who did not meet the KBG diagnostic criteria, 6 were correctly assigned to the missense cluster (Supplemental Table 3B), potentially (in part) driving the observed difference. Finally, the cohort of cases with PTVs was obtained from 1 expert health care center, whereas the missense cohort represents an international collaboration. Larger cohorts are needed to assess whether there are indeed phenotypic differences or whether these results can be explained by cohort effects.

The clinical variability of KBG syndrome is noteworthy, showing considerable phenotypic differences between affected individuals within the same family or between unrelated individuals with the same variant. This variability is best shown by comparing individuals 13 and 14, carrying de novo p.(Glu2522Lys). Although both presented with macrodontia and the characteristic facial appearance, individual 13 exhibited moderate ID, behavioral disturbances, hypotonia, a duplex kidney, strabismus, and normal growth, whereas individual 14 had normal intelligence, no neurobehavioral abnormalities, a submucous cleft palate, moderate hearing loss, mild growth hormone deficiency, and microcephaly. The family with 5 affected individuals also is a key example, in whom 2 presented with macrodontia (individual 21 and 24) and 3 exhibited the characteristic facial gestalt (individuals 20, 23, and 24). We therefore argue not to rule out pathogenicity for individual *ANKRD11* missense variants on inheritance or clinical grounds only. Also for the 2 individuals (individual 3 and 29) not clearly exhibiting symptoms of KBG syndrome, the variants are classified as of uncertain significance and pathogenic when applying American College of Medical Genetics and Genomics/Association for Molecular Pathology criteria (Supplemental Table 5; p.(Leu509Pro) and p.(Leu2605Arg), respectively).[34]

ANKRD11 shows significant regional differences in missense depletion in the general population, with 3 distinct regions of missense tolerance: p.1-p.415 with modest regional missense depletion (o/e = 0.51), p.416-p.2276 tolerating missense variation (o/e = 1.1), and p.2277-p.2664 showing high missense depletion (o/e = 0.11) (unpublished data, Samocha KE, Kosmicki JA, Karczewski KJ, et al. 2017. https://doi.org/10.1101/148353). Consistently, in our cohort, we observed that variants significantly clustered in the highly depleted C-terminal region, which was previously suggested to be implicated in the mechanism underlying KBG syndrome,[13] although we also observed missense variants in the tolerant middle and N-terminal depleted regions. The proportion of independently mutated arginine residues is remarkable (total cohort 12/25 [48%]; RD2, 12/17 [70.6]) and more pronounced than the overrepresentation of mutated arginine residues seen for pathogenic variants underlying genetic disorders in general (15%-20%[35,36]). Arginine is also the most frequently mutated residue in all secondary structures when considering pathogenic variants.[36] Therefore, we hypothesize that the molecular underpinnings of the observed overrepresentation of mutated arginine residues lies in the 3-dimensional structure of ANKRD11, which could not be taken into account, because the crystal structure of ANKRD11 is largely uncharacterized and *ab initio* models are unreliable, despite recent advances in the field.[37] However, on the basis of our *in silico* studies, we argue that if missense variants in *ANKRD11* affect an arginine residue in the C-terminal RD2, it is suggestive for pathogenicity.

Regarding functional impact, most tested missense variants resulted in reduced protein stability, but only for p.(Arg2523Trp), it could be explained by increased proteasome degradation. We hypothesize that variants reducing protein stability without impairment of proteasome degradation affect other mechanisms implicated in protein homeostasis that could be activated by ubiquitination (eg, autophagy). Of note, p.(Arg2523Trp) is located at a putative D-box possibly affecting ANKRD11 ubiquitination and subsequent proteasome degradation. However, the other tested variant at a D-box, p.(Arg2512Gln), showed no impairment of proteasome degradation, which challenges

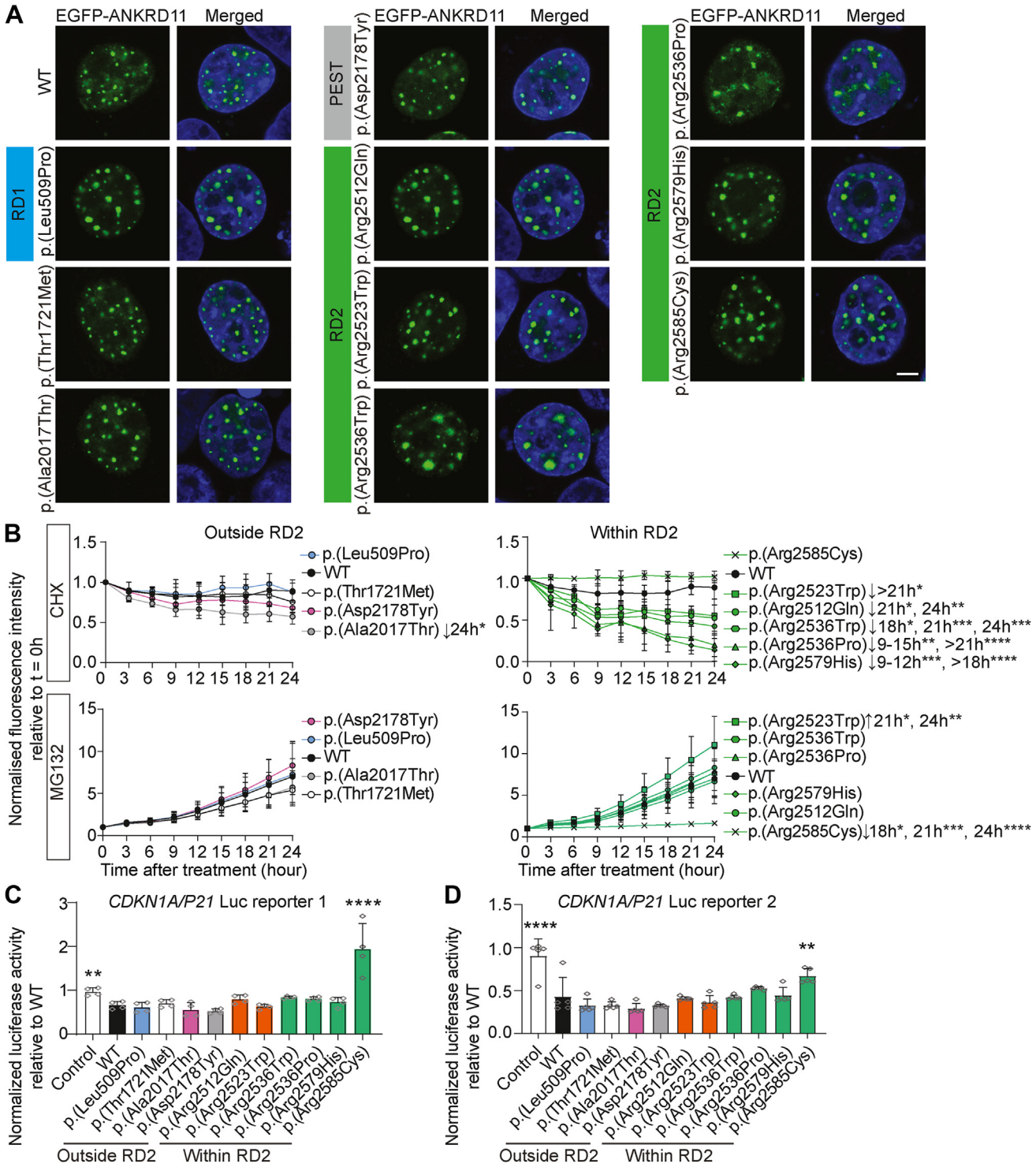

Figure 3 ***ANKRD11* variants in RD2 result in reduced protein stability and impaired proteasome degradation or loss of *CDKN1A/P21* transcriptional repression.** A. Direct fluorescence imaging of cells expressing EGFP-tagged variants of the ANKRD11 protein using confocal microscopy. Wild type and all variants showed a speckle-like pattern in the nucleus. Nuclei are stained with Hoechst 33342 (blue). Protein domains in which variants are located are indicated. Results are representative of 3 independent experiments. Scale bar = 5 μm. B. Relative fluorescence intensity of EGFP-tagged ANKRD11 variants overexpressed in HEK293T/17 cells treated with translation inhibitor cycloheximide (CHX; 50 μg/mL) shown in upper panels and with proteasome inhibitor MG132 (5 μg/mL) in the lower panels. Equal volume of dimethyl sulfoxide was used as a vehicle control. Fluorescence intensity was measured for 24 hours with 3-hour intervals. Values are expressed relative to t = 0 hour and represent the mean ± SD of 3 independent experiments, each preformed in triplicate (*P < .05, **P < .01, ***P < .001, ****P < .0001; 2-way analysis of variance and a post hoc Dunnett's test). C-D. Results of luciferase assay with constructs containing WT and ANKRD11 variants and 2 firefly luciferase reporter constructs with a *CDKN1A/P21* promoter. Values are expressed relative to the control condition that used a EGFP-C2 construct without ANKRD11 and represent the mean ± SD of 4 (C) or 5 (D) independent experiments, each performed in triplicate (*P < .05, **P < .01, ****P < .0001 vs WT; 1-way analysis of variance and a post hoc Dunnett's test). RD, repressor domain; WT, wild type.

the previous suggestion that disruption of the C-terminal D-boxes is the pathophysiological mechanism underlying KBG syndrome.[13] In contrast, the only other variant with altered proteasome degradation, p.(Arg2585Cys), showed reduced proteasome degradation and might slightly increase protein stability, although the latter was not statistically significant. The p.(Arg2585Cys) variant is also the only tested variant resulting in reduced transcriptional repression on *CDKN1A/P21,* consistent with previous observations for p.(Leu2143Val).[14] These findings suggest that p.(Arg2585Cys) results in loss-of-function, despite a potential accumulation of mutant ANKRD11 that contrasts with the dosage reduction seen for the other tested variants. This is further supported by phenotypes of the 2 individuals carrying p.(Arg2585Cys), which are characteristic for KBG syndrome, without apparent differences from phenotypes of individuals with missense variants impairing ANKRD11 stability. Finally, we did not observe changes in ANKRD11 subcellular localization for the assessed missense variants, contrary to what has been reported for Yoda mice with p.(Glu2502Lys).[13]

The 3 variants not showing aberrations in any of our assays are all located outside RD2 and classified as variants of uncertain significance (Supplemental Table 5). Of the 4 tested variants located outside RD2, only p.(Ala2017Thr) affected the assessed protein functions, whereas all tested variants inside RD2 did affect protein function. It is therefore possible that variants outside RD2 exert effects on ANKRD11 functions not captured by our studies. Alternatively, they might alter premessenger RNA splicing (Supplemental Figure 8)[38] despite low SpliceAI scores.[39] On the basis of the role of ANKRD11 in chromatin remodeling, evaluating transcriptomic and epigenetic profiles of individuals or cell-models could increase understanding of the effects of missense variants in the various domains.

In conclusion, our study showed that (de novo) pathogenic missense variants in *ANKRD11* cause the clinically recognizable KBG syndrome with a similar phenotypic spectrum as previously observed for PTVs and microdeletions affecting *ANKRD11*. We showed that loss of transrepression capacity and reduced protein stability are independent molecular mechanisms by which missense variants cause a functional loss of ANKRD11. These findings add to the mechanistic complexity underlying *ANKRD11* haploinsufficiency, already comprising deletion of the locus,[3] putative null alleles,[12] and PTVs escaping the NMD pathway,[7,13] although effects of the latter on protein stability and function have not been elucidated. Because inheritance of pathogenic variants in *ANKRD11* is regularly observed owing to the variability of the associated phenotype, missense variants pose diagnostic challenges, warranting stringent variant classification and careful phenotyping. However, because KBG syndrome is a relatively common cause of genetic NDDs, the involvement of *ANKRD11* missense variants in cohorts of undiagnosed individuals with NDD should be considerable.

Data Availability

Code used for spatial clustering is shared at https://github.com/laurensvdwiel/SpatialClustering. Code used for permutation testing is available at https://github.com/jhampstead/ANKRD11-simulations. All available phenotypic information in Human Phenotype Ontology terminology is shared as a supplement (Supplemental JSON). Model and code for Human Phenotype Ontology-based clustering analysis are available at https://github.com/ldingemans/HPO_clustering_Wang.

Acknowledgments

We are very grateful to all individuals and their families for their participation in this study. This work was financially supported by Aspasia grants of the Dutch Research Council (015.014.036 to T.K. and 015.014.066 to L.E.L.M.V.), Netherlands Organization for Health Research and Development (91718310 to T.K.), and the Max Planck Society (M.M.K.W., S.E.F.). Individual 4 was sequenced at the Scottish Genomes Partnership. The Scottish Genomes Partnership was funded by the Chief Scientist Office of the Scottish Government Health Directorates (SGP/1) and the Medical Research Council Whole Genome Sequencing for Health and Wealth Initiative (MC/PC/15080). The Deciphering Developmental Disorders study presents independent research commissioned by the Health Innovation Challenge Fund (grant number HICF-1009-003). This study makes use of Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (https://www.deciphergenomics.org/), which is funded by Wellcome. See Deciphering Developmental Disorders study[8] or https://www.ddduk.org/access.html for full acknowledgment.

In addition, the collaborations in this study were facilitated by ERN ITHACA, one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States, cofunded by European Commission. The aims of this study contribute to the Solve-RD project (E.d.B., A.S.D.P., L.F., C.G., T.K., A.V., L.E.L.M.V.), which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 779257.

Author Information

Conceptualization: C.W.O., M.M.K.W., T.K.; Data Curation: E.d.B., C.W.O., D.R., T.A., R.B., M.B.-H., B.A., N.C., A.-S.D.-P., O.D., C.D., F.E., H.Z.E., L.F., S.F.B., D.G., J.A.C.G., B.M.H., U.K., A.L.-A., G.L., S.A.L., I.M.J.M., R.M.G., K.G.M., S.O., R.P., A.P., J.v.R., G.W.E.S., E.S., A.S., P.J.v.d.S., A.P.A.S., S.M.A.S., I.V., E.V.-D., A.V., S.M.W., M.W., K.J.L., T.K.; Formal Analysis: E.d.B.,

M.M.K.W., C.W.O., L.E.L.M.V., S.E.F., T.K.; Investigation: R.A.K., E.d.B, M.M.K.W., C.W.O., T.K.; Software: L.L., A.J.M.D., J.E.H., C.G.; Writing-original draft: E.d.B., M.M.K.W., C.W.O., T.K.; Supervision: C.W.O., M.M.K.W., T.K.; All authors contributed to the final version of the manuscript.

Ethics Declaration

We obtained informed consent to publish unidentifiable data for all individuals reported in this study. Specific consent was obtained for publication of clinical photographs. Consent procedures were in accordance with the Declaration of Helsinki and local ethical guidelines of the participating centers. The institutional review board 'Commissie Mensgebonden Onderzoek Regio Arnhem-Nijmegen' approved this study under number 2011/188 and 2022-13611. Number 2011/188 refers to performing diagnostic exome sequencing. Discovery of novel syndromes and description of clinical cohorts from this series can be taken as such. Number 2022-13611 refers to publishing Human Phenotype Ontology data for individuals in Biobank Genetics and Rare Diseases and Biobank Intellectual Disability of the Radboudumc. All the appropriate institutional forms have been archived locally.

Conflict of Interest

H.Z.E. and K.G.M. are employees of GeneDx, Inc. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2022.06.007) contains supplementary material, which is available to authorized users.

Authors

Elke de Boer[1,2], Charlotte W. Ockeloen[1,*], Rosalie A. Kampen[3], Juliet E. Hampstead[1,4], Alexander J.M. Dingemans[1,2], Dmitrijs Rots[1,2], Lukas Lütje[3], Tazeen Ashraf[5,6], Rachel Baker[7], Mouna Barat-Houari[8], Brad Angle[7], Nicolas Chatron[9,10], Anne-Sophie Denommé-Pichon[11,12], Orrin Devinsky[13], Christèle Dubourg[14,15], Frances Elmslie[16], Houda Zghal Elloumi[17], Laurence Faivre[11,18,19], Sarah Fitzgerald-Butt[20], David Geneviève[21], Jacqueline A.C. Goos[22,23], Benjamin M. Helm[20,24], Usha Kini[25], Amaia Lasa-Aranzasti[26], Gaetan Lesca[9,10], Sally A. Lynch[27], Irene M.J. Mathijssen[22], Ruth McGowan[28], Kristin G. Monaghan[17], Sylvie Odent[29], Rolph Pfundt[1], Audrey Putoux[30,31], Jeroen van Reeuwijk[1,2], Gijs W.E. Santen[32], Erina Sasaki[25], Arthur Sorlin[11,18], Peter J. van der Spek[23], Alexander P.A. Stegmann[1,33], Sigrid M.A. Swagemakers[23], Irene Valenzuela[26], Eléonore Viora-Dupont[18], Antonio Vitobello[11,12], Stephanie M. Ware[20,34], Mathys Wéber[18], Christian Gilissen[1,4], Karen J. Low[35], Simon E. Fisher[2,3], Lisenka E.L.M. Vissers[1,2], Maggie M.K. Wong[3], Tjitske Kleefstra[1,2,36]

Affiliations

[1]Department of Human Genetics, Radboudumc, Nijmegen, The Netherlands; [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; [3]Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [4]Radboud Institute for Molecular Life Sciences, Radboudumc, Nijmegen, The Netherlands; [5]Department of Clinical Genetics, Great Ormond Street Hospital for Children NHS Foundation Trust, London, United Kingdom; [6]Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom; [7]Advocate Children's Hospital, Park Ridge, IL; [8]Genetic Laboratory of Rare and Autoinflammatory Diseases, Department of Medical Genetics, Rare Diseases and Personalized Medicine, Centre Hospitalier Universitaire de Montpellier, Montpellier, France; [9]Service de Génétique, Hospices Civils de Lyon, Bron, France; [10]Institut NeuroMyoGene, CNRS UMR5310, INSERM U1217, Université Claude Bernard Lyon 1, Lyon, France; [11]Génétique des Anomalies du Développement, Université de Bourgogne Franche-Comté, UMR1231-Inserm, Dijon, France; [12]Laboratoire de Génétique Chromosomique et Moléculaire, UF6254 Innovation en Diagnostic Génomique des Maladies Rares, Centre Hospitalier Universitaire de Dijon, Dijon, France; [13]Department of Neurology, NYU Grossman School of Medicine, NYU Langone Health, New York, NY; [14]Service de Génétique Moléculaire et Génomique Médicale, CHU de Rennes, Rennes, France; [15]University of Rennes, CNRS, IGDR, UMR 6290, Rennes, France; [16]South West Thames Regional Clinical Genetics Service, St George's Hospital, University of London, London, United Kingdom; [17]GeneDx, Gaithersburg, MD; [18]Centre de Génétique et Centre de Référence Anomalies du Développement et Syndromes Malformatifs de l'Interrégion Est, Centre Hospitalier Universitaire Dijon, Dijon, France; [19]Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (TRANSLAD), Centre Hospitalier Universitaire Dijon, Dijon, France; [20]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indiana University, Indianapolis, IN; [21]Medical Genetic Department, Rare Diseases and Personalized Medicine, Montpellier University, Inserm U1183, CHU

Montpellier, Montpellier, France; [22]Department of Plastic and Reconstructive Surgery and Hand Surgery, Dutch Craniofacial Center, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [23]Department of Bioinformatics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [24]Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN; [25]Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom; [26]Department of Clinical and Molecular Genetics, Vall d'Hebron University Hospital and Medicine Genetics Group, Vall d'Hebron Research Institute, Barcelona, Spain; [27]Department of Clinical Genetics, Children's Health Ireland at Crumlin and Temple Street, Dublin, Ireland; [28]West of Scotland Centre for Genomic Medicine, Queen Elizabeth University Hospital, Scottish Genomes Partnership, Glasgow, United Kingdom; [29]CHU Rennes, Service de Génétique Clinique, Centre de Référence Maladies Rares CLAD-Ouest, ERN ITHACA, Hôpital Sud, Rennes, France; [30]Service de Génétique - Centre de Référence Anomalies du Développement, Hospices Civils de Lyon, Bron, France; [31]Équipe GENDEV, Centre de Recherche en Neurosciences de Lyon, INSERM U1028 CNRS UMR5292, Université Claude Bernard Lyon 1, Lyon, France; [32]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands; [33]Department of Clinical Genetics, Maastricht University Medical Center+, Maastricht University, Maastricht, The Netherlands; [34]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN; [35]Department of Clinical Genetics, University Hospital Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom; [36]Center of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands

References

1. Herrmann J, Pallister PD, Tiddy W, Opitz JM. The KBG syndrome-a syndrome of short stature, characteristic facies, mental retardation, macrodontia and skeletal anomalies. *Birth Defects Orig Artic Ser*. 1975;11(5):7–18.
2. Ockeloen CW, Willemsen MH, de Munnik S, et al. Further delineation of the KBG syndrome caused by ANKRD11 aberrations. *Eur J Hum Genet*. 2015;23(9):1176–1185. Published correction appears in *Eur J Hum Genet*. 2015;23(9):1270. https://doi.org/10.1038/ejhg.2014.253
3. Goldenberg A, Riccardi F, Tessier A, et al. Clinical and molecular findings in 39 patients with KBG syndrome caused by deletion or mutation of ANKRD11. *Am J Med Genet A*. 2016;170(11):2847–2859. http://doi.org/10.1002/ajmg.a.37878.
4. Low K, Ashraf T, Canham N, et al. Clinical and genetic aspects of KBG syndrome. *Am J Med Genet A*. 2016;170(11):2835–2846. http://doi.org/10.1002/ajmg.a.37842.
5. Gnazzo M, Lepri FR, Dentici ML, et al. KBG syndrome: common and uncommon clinical features based on 31 new patients. *Am J Med Genet A*. 2020;182(5):1073–1083. http://doi.org/10.1002/ajmg.a.61524.
6. Skjei KL, Martin MM, Slavotinek AM. KBG syndrome: report of twins, neurological characteristics, and delineation of diagnostic criteria. *Am J Med Genet A*. 2007;143A(3):292–300. http://doi.org/10.1002/ajmg.a.31597.
7. Sirmaci A, Spiliopoulos M, Brancati F, et al. Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am J Hum Genet*. 2011;89(2):289–294. http://doi.org/10.1016/j.ajhg.2011.06.007.
8. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. 2015;519(7542):223–228. http://doi.org/10.1038/nature14135.
9. Kaplanis J, Samocha KE, Wiel L, et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*. 2020;586(7831):757–762. http://doi.org/10.1038/s41586-020-2832-5.
10. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434–443. Published correction appears in *Nature*. 2021;590(7846):E53. Published correction appears in *Nature*. 2021;597(7874):E3-E4. https://doi.org/10.1038/s41586-020-2308-7
11. Morel Swols D, Foster J 2nd, Tekin M. KBG syndrome. *Orphanet J Rare Dis*. 2017;12(1):183. http://doi.org/10.1186/s13023-017-0736-8.
12. Cucco F, Sarogni P, Rossato S, et al. Pathogenic variants in EP300 and ANKRD11 in patients with phenotypes overlapping Cornelia de Lange syndrome. *Am J Med Genet A*. 2020;182(7):1690–1696. http://doi.org/10.1002/ajmg.a.61611.
13. Walz K, Cohen D, Neilsen PM, et al. Characterization of ANKRD11 mutations in humans and mice related to KBG syndrome. *Hum Genet*. 2015;134(2):181–190. http://doi.org/10.1007/s00439-014-1509-2.
14. Zhang T, Yang Y, Yin X, et al. Two loss-of-function ANKRD11 variants in Chinese patients with short stature and a possible molecular pathway. *Am J Med Genet A*. 2021;185(3):710–718. http://doi.org/10.1002/ajmg.a.62024.
15. Gallagher D, Voronova A, Zander MA, et al. Ankrd11 is a chromatin regulator involved in autism that is essential for neural development. *Dev Cell*. 2015;32(1):31–42. http://doi.org/10.1016/j.devcel.2014.11.031.
16. Zhang A, Yeung PL, Li CW, et al. Identification of a novel family of ankyrin repeats containing cofactors for p160 nuclear receptor coactivators. *J Biol Chem*. 2004;279(32):33799–33805. http://doi.org/10.1074/jbc.M403997200.
17. Zhang A, Li CW, Chen JD. Characterization of transcriptional regulatory domains of ankyrin repeat cofactor-1. *Biochem Biophys Res Commun*. 2007;358(4):1034–1040. http://doi.org/10.1016/j.bbrc.2007.05.017.
18. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062–D1067. http://doi.org/10.1093/nar/gkx1153.
19. Hanly C, Shah H, Au PYB, Murias K. Description of neurodevelopmental phenotypes associated with 10 genetic neurodevelopmental disorders: A scoping review. *Clin Genet*. 2021;99(3):335–346. http://doi.org/10.1111/cge.13882.
20. Barbaric I, Perry MJ, Dear TN, et al. An ENU-induced mutation in the Ankrd11 gene results in an osteopenia-like phenotype in the mouse mutant Yoda. *Physiol Genomics*. 2008;32(3):311–321. http://doi.org/10.1152/physiolgenomics.00116.2007.
21. Köhler S, Carmody L, Vasilevsky N, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res*. 2019;47(D1):D1018–D1027. http://doi.org/10.1093/nar/gky1105.
22. den Hoed J, de Boer E, Voisin N, et al. Mutation-specific pathophysiological mechanisms define different neurodevelopmental disorders associated with SATB1 dysfunction. *Am J Hum Genet*. 2021;108(2):346–356. http://doi.org/10.1016/j.ajhg.2021.01.007.
23. Deng Y, Gao L, Wang B, Guo X. HPOSim: an R package for phenotypic similarity measure and enrichment analysis based on the human phenotype ontology. *PLoS One*. 2015;10(2):e0115692. http://doi.org/10.1371/journal.pone.0115692.
24. Wang JZ, Du Z, Payattakool R, Yu PS, Chen CF. A new method to measure the semantic similarity of GO terms. *Bioinformatics*. 2007;23(10):1274–1281. http://doi.org/10.1093/bioinformatics/btm087.

25. Lelieveld SH, Wiel L, Venselaar H, et al. Spatial clustering of de novo missense mutations identifies candidate neurodevelopmental disorder-associated genes. *Am J Hum Genet*. 2017;101(3):478–484. http://doi.org/10.1016/j.ajhg.2017.08.004.
26. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. http://doi.org/10.1186/s13059-016-0974-4.
27. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. *Genetics*. 2000;156(1):297–304. http://doi.org/10.1093/genetics/156.1.297.
28. Nishio H, Walsh MJ. CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. *Proc Natl Acad Sci U S A*. 2004;101(31):11257–11262. http://doi.org/10.1073/pnas.0401343101.
29. el-Deiry WS, Tokino T, Velculescu VE, et al. WAF1, a potential mediator of p53 tumor suppression. *Cell*. 1993;75(4):817–825. http://doi.org/10.1016/0092-8674(93)90500-p.
30. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36(10):928–930. http://doi.org/10.1002/humu.22844.
31. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am J Hum Genet*. 2009;84(4):524–533. http://doi.org/10.1016/j.ajhg.2009.03.010.
32. Jehl P, Manguy J, Shields DC, Higgins DG, Davey NE. ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences. *Nucleic Acids Res*. 2016;44(W1):W11–W15. http://doi.org/10.1093/nar/gkw265.
33. Kaufman L, Rousseeuw PJ. Clustering by means of medoids. In: *Proceedings of the Statistical Data Analysis Based on the L1 Norm Conference August 31, 1987 (Vol. 31)*. https://wis.kuleuven.be/stat/robust/papers/publications-1987/kaufmanrousseeuw-clusteringbymedoids-l1norm-1987.pdf
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–424. http://doi.org/10.1038/gim.2015.30.
35. Schulze KV, Hanchard NA, Wangler MF. Biases in arginine codon usage correlate with genetic disease risk. *Genet Med*. 2020;22(8):1407–1412. http://doi.org/10.1038/s41436-020-0813-6.
36. Khan S, Vihinen M. Spectrum of disease-causing mutations in protein secondary structures. *BMC Struct Biol*. 2007;7:56. http://doi.org/10.1186/1472-6807-7-56.
37. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589. http://doi.org/10.1038/s41586-021-03819-2.
38. Fadaie Z, Khan M, Del Pozo-Valero M, et al. Identification of splice defects due to noncanonical splice site or deep-intronic variants in ABCA4. *Hum Mutat*. 2019;40(12):2365–2376. http://doi.org/10.1002/humu.23890.
39. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting splicing from primary sequence with deep learning. *Cell*. 2019;176(3):535–548.e24. http://doi.org/10.1016/j.cell.2018.12.015.