The *RBMX* gene as a Candidate for the Shashi X-linked Intellectual Disability Syndrome (SMRXS)

Vandana Shashi[1], Pingxing Xie[2], Kelly Schoch[1], David B. Goldstein[2], Timothy D. Howard[3], Margaret N. Berry[3], Charles E. Schwartz[4], Kenneth Cronin[2], Samantha Sliwa[2], Andrew Allen[5], Anna C. Need[6]

[1]Department of Pediatrics, Division of Medical Genetics, Duke University Medical Sciences, Durham NC

[2]Center for Human Genome Variation, Duke University School of Medicine

[3]Department of Pediatrics, Wake Forest University Health Sciences, Winston-Salem NC

[4]Greenwood Genetic Center, Greenwood SC

[5]Department of Biostatistics, Duke University Health Sciences

[6]Division of Brain Sciences, Imperial College, London, UK

Corresponding author:

Vandana Shashi, MD

Box 103857,

Duke University Health Sciences,

Durham NC 27710

Tel: (919) 681-2616

Fax: (919) 668-0414

Email: vandana.shashi@duke.edu

No conflicts of interest

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cge.12511

Abstract

A novel X-linked intellectual disability (XLID) syndrome with moderate intellectual disability and distinguishing craniofacial dysmorphisms had been previously mapped to the Xq26-q27 interval. On whole exome sequencing in the large family originally reported with this disorder, we identified a 23 bp frameshift deletion in the RNA Binding Motif Protein X-linked (*RBMX*) gene at Xq26 in the affected males (n=7), one carrier female, absent in unaffected males (n=2) and in control databases (7800 exomes). The *RBMX* gene has not been previously causal of human disease. We examined the genic intolerance scores for the coding regions and the non-coding regions of *RBMX*; the findings were indicative of *RBMX* being relatively intolerant to loss of function variants, a distinctive pattern seen in a subset of XLID genes. Prior expression and animal modeling studies indicate that loss of function of *RBMX* results in abnormal brain development. Our finding putatively adds a novel gene to the loci associated with XLID and may enable the identification of other individuals affected with this distinctive syndrome.

Introduction

A concentration of genes associated with intelligence on the X-chromosome confers males with a higher vulnerability to X-linked recessive intellectual disability (XLID) (1, 2). Causal mutations in XLID genes are usually loss-of-function, impairing inhibitory and excitatory neurotransmission (3). Approximately 200 XLID disorders have been described, with over 100 genes being identified so far (http://www.ggc.org/research/molecular-studies/xlid.html) (4). Whole exome sequencing (WES) has accelerated the discovery of these genes and has also resulted in the detection of spurious gene associations, with approximately 10% being discovered as being erroneous, enabling more accurate etiological associations (5). Recent advances such as the prioritization of genes based on intolerance to functional variants in the coding (RVIS) (6) and non-coding regions (ncRVIS) (manuscript in preparation) can enable further refinement of the judgement of whether a gene could be associated with XLID, since genes associated with Mendelian diseases have been generally found to be intolerant to functional variants (6).

We had previously reported a large family in North Carolina with a unique form of XLID with moderate intellectual disability and craniofacial dysmorphisms; the gene was mapped to Xq26-q27 (7). Subsequently, Castro et al.,(8) reported two brothers who they believed had the same condition based on clinical features and haplotyping. However, the underlying gene responsible for this unique XLID syndrome (OMIM%300238, Shashi X-linked Mental Retardation Syndrome, SMRSX) remained unknown. Recently, we contacted the original family members again to identify the underlying gene with WES.

Methods

Subjects: Six of the seven affected males in the original family that we had reported with SMRXS were reevaluated (7); affected male III-2 is now deceased (Figure 1). A seventh affected male (VI-2) in this family was born subsequent to the original report and is now included. The ages of the affected males ranged from 13-65 years. One carrier female (IV-11) in the family was also enrolled. Details of clinical evaluations of the subjects are available in the previous publication (7)

The study was approved by the institutional review boards of Wake Forest Health Sciences and Duke University Health Sciences.

WES: Previously extracted peripheral blood DNA samples and/or new DNA from the participants were utilized. WES was performed on two affected subjects (IV-4 and IV-5), using previously described methods (9, 10). Control cohorts consisted of subjects from Duke (n=1300) and from the NHLBI GO Exome Sequencing Project (ESP, n=6503) (http://evs.gs.washington.edu/EVS/). None of the controls had been ascertained for severe rare genetic disorders or intellectual disabilities.

We searched for hemizygous X chromosome variants (with special attention to the Xq26-q27 interval) that were shared between IV-4 and IV-5, but were not carried by any male controls, focusing on nonsynonymous variants, variants in essential splice sites, and small insertion/deletion variants located in the protein-coding regions. The variant of interest was Sanger sequenced on all participants from the family using Applied Biosystems 3730 DNA Analyzer (Foster City, CA), Further confirmation of the variant was performed in a CLIA-certified laboratory (GeneDx, Gaithersburg, MD) on one affected male (IV-3) to communicate the results to the family. Sanger sequencing of the RBMX gene was performed on the two affected males reported by Castro et al (8).

RVIS) and ncRVIS Scores: We used the RVIS (6) and ncRVIS to predict how likely mutations in the *RBMX* gene are selected against in the healthy human population. The

RVIS score is based on protein coding region variants and the ncRVIS score on the 5' and 3' UTRs and the 250bp upstream of the gene's transcription start site

Results

Subjects: The six affected males now range from 28 to 65 years. All have the distinctive coarse faces, prominent supraorbital ridges, narrow palpebral fissures, bulbous nose, prominent lower lip, large ears and obesity (Figure 2) (7). Affected males IV-3, IV-4, IV-5 and IV-6 were diagnosed with bilateral sensorineural hearing loss in their early 50s. All the affected males live at home with their guardians. The seventh affected male (III-2), died in his 80s, due to age related complications. Male VI-2vis 13 years-old, has moderate intellectual disability and epilepsy. The PI had examined him during early childhood and noted the characteristic facial features and obesity. Due to aggressiveness he could not be reevaluated.

WES: Exome sequencing data was of high quality; average coverage of the X chromosome was 44-fold for IV-4 and 57-fold for IV-5, with 5X coverage for 91.8% and 96.1% of targeted regions respectively. Three hemizygous variants on the X chromosome were shared by IV-4 and IV-5, absent in all internal male controls and in the NHLBI male subjects. The first was a missense variant in *ARSF* located in the Xp22.33 region (p.V170I, NP_004033.2), predicted as benign by PolyPhen-2 (11). The *ARSF* gene is outside the linkage interval and was thus not pursued further. The second was a missense mutation in *UTP14A* (p.R647C, NP_006640.2), is in the Xq26.1 interval (chrX: 129,040,096-129,063,737), also predicted as benign by PolyPhen-2. In the NHLBI EVS database 14 other types of missense variants in *UTP14A* were seen in males. The X-chromosome RVIS score for UTP14A is on the 80.3 percentile with the ncRVIS score on the 37th percentile, making it a gene tolerant to functional variants. Additionally, the nature of the mutation (missense) and the presence of other missense variants in healthy control males resulted in this variant not being pursued further.

The third variant was a 23-bp deletion causing a frameshift in exon 9 of *RBMX* (Figure 3). It changes the 336 position glutamic acid to glycine, deletes the following seven amino acids, and generates a stop codon after another seven amino acids. It was not observed in any of the 1300 Duke controls or 6503 NHLBI ESP subjects. We looked for other indels and damaging mutations in *RBMX* in both control cohorts, and only one in-frame deletion was observed (chrX: 135954483_135954485_DEL_ATA, located in exon 10, only in the 3rd isoform of *RBMX.* There was no experimental confirmation for isoform 3 based on UniProt database. Sanger sequencing demonstrated the frameshift deletion in all seven affected males (IV-3, IV-4, IV-5, IV-6, V-3, V-4 and VI-2), one carrier female (IV-11) and was absent in two normal males (IV-7, IV-10). However, Sanger sequencing on the two affected males from the Castro et al report (8) did not show the 23 bp deletion or other variants.

RVIS and ncRVIS scores: The RVIS score for *RBMX* was on the 73rd percentile genome-wide and on the 74.33 percentile for the X-chromosome genes suggesting that it may be tolerant of functional mutations. We then examined the preferential intolerance of *RBMX* to nonsense mutations: across the ESP6500 controls, there were 16 observed variants (none was nonsense) and with the Broad institute Z-score for loss of function mutations in *RBMX* being 0.8, there was insufficient resolution to assess preferential intolerance of nonsense mutations in *RBMX*. Additionally, there is a lack of methods to assess preferential depletion of frameshift mutations within any gene. These factors limited our analyses to accurately assess the tolerance to loss of function mutations in the coding regions of *RBMX*. However, the ncRVIS score for *RBMX* places it around the 15th percentile (genome-wide) and on the 18.42 percentile for the X chromosome providing indirect evidence that it may indeed be intolerant to loss of function mutations.

Discussion

The absence of deleterious variants in *RBMX* in over 7800 control exomes suggests that protein disrupting mutations in this gene are not compatible with normal cognitive function. The importance of controls that are representative of the 'general population' in studies of XLID is highlighted by a report that found that at least 10% of the 100 genes that had been deemed to cause XLID were incorrect; the authors recommended that all putative genes for XLID be validated against large databases such as the NHLBI EVS (5). Further supporting evidence for *RBMX* as a candidate gene for SRMXS is provided by the highly conserved nature of this gene (12), the type of mutation and the location of *RBMX* within the previously linked interval in the Xq26 region. *RBMX*'s function as an RNA binding protein would make it also a suitable candidate gene, since other RNA binding proteins such as the FMRP in Fragile X syndrome have been reported to account for 3-5% of XLID conditions (2).

To further refine the possible causal role of *RBMX*, we pursued two cutting-edge tools that rank the tolerance of genes to functional variants: RVIS and ncRVIS. In 86 syndromic XLID genes from an OMIM "X-linked mental retardation' key match, the RVIS scores ranged from 0.4th percentile (*HUWE1*) to 95.7th percentile (ZNF81), with *FMR1* (Fragile X syndrome) on the 70.3rd percentile. Interestingly, *OPHN1*, causing XLID with cerebellar hypoplasia (13), has a high RVIS score (80.2 percentile), but an ncRVIS score at 2.7th percentile, raising the possibility that the mechanism of disease is loss of function. Indeed, all the reported disease causing mutations in *OPHN1* are loss of function mutations (14). Similarly, *FMR1* has a low ncRVIS score of 17th percentile and over 80% of reported mutations cause loss of function. With an 18th percentile ncRVIS score, *RBMX* may be part of a subset of XLID genes that follow a pattern of high RVIS scores and low ncRVIS scores, with disease causing mutations being predominantly loss of function.

The widespread expression *of RBMX* may be indicative of its effects on the brain as well as other general developmental processes. An orthologue for the human *RBMX* gene in zebra fish showed expression in the later stages of brain development; in knockouts of *RBMX* in

zebra fish embryos the smaller forebrain could be rescued by injection of wild type *RBMX* (15). Abnormalities in neurogenesis, neural plate and neural crest development were seen in a *RBMX* knockdown model of the African clawed frog (16). Thus, these animal studies support the possibility that *RBMX* is involved in neurodevelopment. Possible mechanisms such as altered cellular localization of this RNA binding protein, or a toxic effect due the unstable truncated protein causing the abnormal neurodevelopment could be the topic of further research.

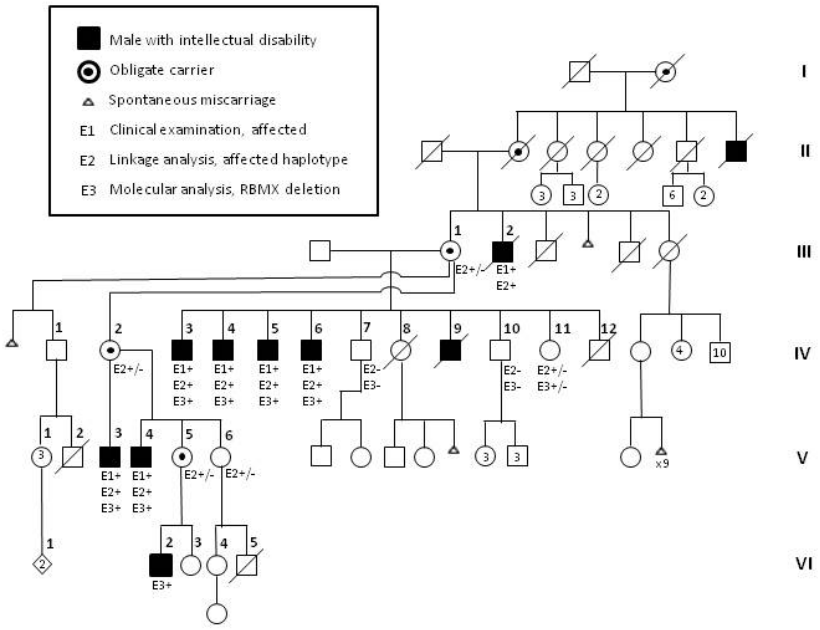
The prevalence of SMRSX remains unclear. Two brothers with XLID were reported as having SMRSX several years ago by Castro et al (8). We performed Sanger sequencing on these two brothers and did not find the 23 bp deletion or other variants in *RBMX.* However, the interval shared by the two affected brothers was significantly larger, from Xq21.1-Xq27. Additionally, the salient features of SMRSX, including the facial cephalometric findings and the facial gestalt were not evident in the two brothers. Thus, the family reported by Castro et al is likely to have another form of XLID. Clark-Baraitser and Atkins type XLID (17, 18) have moderate phenotypic overlap with SMRXS (7) and the genes remain unmapped, but we were unsuccessful in obtaining DNA from these cohorts to look for *RBMX* variants.

We acknowledge that the RBMX mutation could be segregating with the linkage interval and that the causal mutation may be in another gene/intron that was not detected by WES. We also did not pursue functional studies. However, since the frameshift mutation would cause decreased expression, demonstrating this would not provide any further evidence that the mutation is causal. The ncRVIS scoring system is relatively new and has not been widely validated yet. Despite these weaknesses, the absence of the frameshift deletion and other deleterious mutations in *RBMX* in a large number of control exomes, the pattern of RVIS and ncRVIS scores for RBMX being consistent with a pattern seen in a subset of X-linked disease causing genes, the biological function of this gene, its highly conserved status (12) and the findings of abnormal neurodevelopment in animal *RBMX* knockout models all suggest that this is a reasonable candidate gene for this unique XLID syndrome.

Bibliography

1. Gecz J, Shoubridge C, Corbett M. The genetic landscape of intellectual disability arising from chromosome X. Trends Genet 2009: 25: 308-316.
2. Lubs HA, Stevenson RE, Schwartz CE. Fragile X and X-linked intellectual disability: four decades of discovery. Am J Hum Genet 2012: 90: 579-590.
3. Bassani S, Zapata J, Gerosa L et al. The neurobiology of X-linked intellectual disability. Neuroscientist 2013: 19: 541-552.
4. Tarpey PS, Smith R, Pleasance E et al. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. Nat Genet 2009: 41: 535-543.
5. Piton A, Redin C, Mandel JL. XLID-Causing Mutations and Associated Genes Challenged in Light of Data From Large-Scale Human Exome Sequencing. Am J Hum Genet 2013: 93: 368-383.
6. Petrovski S, Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. . Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes. . PLoS Genet 2013: 9, e1003709.
7. Shashi V, Berry MN, Shoaf S et al. A unique form of mental retardation with a distinctive phenotype maps to Xq26-q27. Am J Hum Genet 2000: 66: 469-479.
8. Castro NH, dos Santos RC, Nelson R et al. Shashi XLMR syndrome: report of a second family. Am J Med Genet A 2003: 118A: 49-51.
9. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010: 26: 589-595.
10. Cingolani P, Platts A, Wang le L et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012: 6: 80-92.
11. Adzhubei IA, Schmidt S, Peshkin L et al. A method and server for predicting damaging missense mutations. Nat Methods 2010: 7: 248-249.
12. Lingenfelter PA, Delbridge ML, Thomas S et al. Expression and conservation of processed copies of the RBMX gene. Mamm Genome 2001: 12: 538-545.
13. Billuart P, Bienvenu T, Ronce N et al. Oligophrenin 1 encodes a rho-GAP protein involved in X-linked mental retardation. Pathol Biol (Paris) 1998: 46: 678.
14. https://www.ncbi.nlm.nih.gov/clinvar/. Clinvar.
15. Tsend-Ayush E, O'Sullivan LA, Grutzner FS et al. RBMX gene is essential for brain development in zebrafish. Dev Dyn 2005: 234: 682-688.
16. Dichmann DS, Fletcher RB, Harland RM. Expression cloning in Xenopus identifies RNA-binding proteins as regulators of embryogenesis and Rbmx as necessary for neural and muscle development. Dev Dyn 2008: 237: 1755-1766.
17. Atkin JF, Flaitz K, Patil S et al. A new X-linked mental retardation syndrome. Am J Med Genet 1985: 21: 697-705.
18. Baraitser M, Reardon W, Vijeratnam S. Nonspecific X-linked mental retardation with macrocephaly and obesity: a further family. Am J Med Genet 1995: 57: 380-384.

Figure 1: Updated pedigree of large family reported with SMRSX. Since the time of the initial publication another affected male (VI-2) has been identified and affected male, III-2 has died in his 80s due to age related complications.



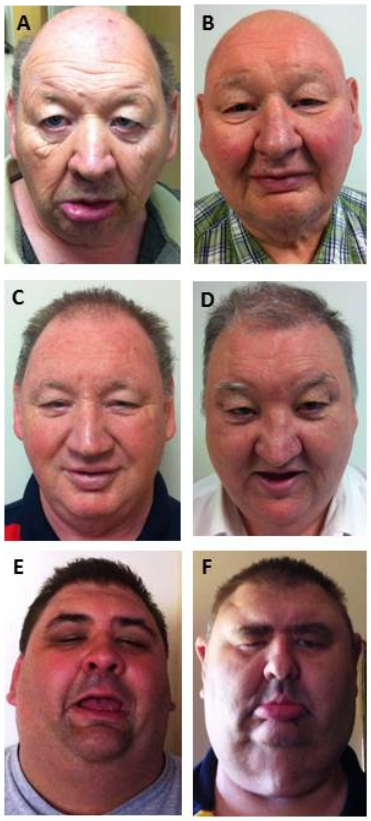


Figure 2: Composite picture of affected adult males, including IV-3 (A), IV-4 (B), IV-5 (C), IV-6 (D), V-3 (E), and V-4 (F). All six affected males continue to show the characteristic facial features of coarseness, prominent supraorbital ridges, periorbital fullness, bulbous nose, prominent lower lip and large ears

Figure 3: Illustration of the Sanger sequencing results of the frameshift deletion in *RBMX* (chrX:135956417_135956439_DEL_CATAGAAGGGGGAAGCCCTCTTT) in selected family members with SMRXS. The 23 bp deletion is seen in the two affected males IV-4 (A), IV-5 (B) and the carrier female IV-11 (C), but not in the unaffected male IV-10 (D).