Meta Gene 13 (2017) 57–62

Contents lists available at ScienceDirect

Meta Gene

journal homepage: www.elsevier.com/locate/mgene

PTBP1: A candidate gene for intellectual disability

Asem M. Alkhateeb[a,*], Wesal Habbab[b], Samah K. Aburahma[c], I. Richard Thompson[b]

[a] Biotechnology and genetics department, Jordan University of Science and Technology, Irbid 22110, P.O. Box 3030, Jordan
[b] Qatar Biomedical Research Institute, Hamad bin Khalifa University, Doha, P.O. Box 5825, Qatar
[c] Neurosciences department, Jordan University of Science and Technology, Irbid 22110, P.O. Box 3030, Jordan

ARTICLE INFO

Keywords:
Intellectual disability
Exome
PTBP1
Synonymous

ABSTRACT

Intellectual disability (ID) is a common heterogeneous disease. Many genes have been implicated in the etiology of ID, yet gene discovery continues at a fast pace. In this study, we investigated the genetic cause of ID in a consanguineous Arab family. For that purpose we utilized autozygosity mapping, candidate gene testing, and whole exome sequencing to find the causative mutation. We identified a synonymous mutation, c.708C > T, in the polypyrimidine tract binding protein 1 (*PTBP1*) gene that fully segregated with the ID phenotype; Mutationtaster predicted the mutation to be disease causing. Analysis of the mutation revealed its effect on an exonic splice enhancer suggesting a change in the potential binding to its target sequence. This mutation was not found in 514 ethnically matched chromosomes and only one allele harbored the same mutation out of 120,782 alleles in the Exome Aggregation Consortium (ExAC) database. PTBP1 plays a major role in splicing related to neuronal differentiation; its repressive activity regulates activation of a brain specific alternative splicing program. This is the first study to implicate a *PTBP1* mutation in human disease.

1. Introduction

Intellectual disability (ID) is a disorder with onset during the developmental period, before the age of 18, which includes both adaptive and intellectual functioning deficits in conceptual, practical, and social domains (American Psychiatric Association, 2013). ID is a common disease with a prevalence of 1.5–2% (Delobel-Ayoub et al., 2015), and is expected to be much higher in highly inbred populations (Iqbal and van Bokhoven, 2014; Saad et al., 2014). More than 700 hundred genes have been implicated in ID (Vissers et al., 2016), around 100 of which are X-linked, and while it is believed that most X-linked genes have been identified already, it is likely that more autosomal genes are still to be identified. One could think that most important or frequent ID genes have already been discovered but two recently identified genes, *ARID1B* and *DDX3X*, may each cause disease in > 1% of ID patients (Santen et al., 2012, Snijders Blok et al., 2015) which suggests that further work is warranted to profile all ID genes. Although autosomal recessive mutations may not be a major etiological factor in the genetics of ID in outbred populations, it does play a key role in inbred populations (Musante and Ropers, 2014; Corbett et al., 2010; Brečević et al., 2015).

Abnormality in diverse sets of proteins causes ID which points to a critical and non-redundant function of those proteins in the regular pathways leading to normal intellect. Post-transcriptional control of RNA processing within neurons provides spatial and temporal regulation of gene expression critical for proper brain development. Control of RNA processing steps is strictly controlled by RNA binding proteins. Mutations of multiple RNA binding proteins are found to cause many human neurological and neurodevelopmental diseases (Darbelli and Richard, 2016; Lee et al., 2016; McMahon et al., 2016). Since many mutations are causing human disease through a direct effect on mRNA splicing, polymorphisms found in splicing regulatory genes can play key roles in altering disease status and severity in affected individuals (Vuong et al., 2016). Splicing regulatory mechanisms can affect a wide range of neuronal development, from neurogenesis to mature synaptic function.

In this study we investigated the genetic etiology of ID by SNP genotyping and whole exome sequencing in a consanguineous family that harbors two affected female siblings. We found an exceptionally rare homozygous co-segregating mutation in *PTBP1*. Polypyrimidine tract binding protein 1 (*PTBP1*) is an RNA binding protein that plays a key role in neurogenesis and brain-specific alternative splicing (Gueroussov et al., 2015). Additionally, it regulates alternative splicing of micro-exons in human brain transcripts (Li et al., 2015). Alternating levels of *PTBP1* expression affect neuronal fate and early neuronal differentiation. Its function makes it a plausible candidate gene for affecting the intellectual development in humans.

* Corresponding author.
E-mail address: asemalkhateeb@just.edu.jo (A.M. Alkhateeb).

http://dx.doi.org/10.1016/j.mgene.2017.05.002
Received 8 February 2017; Received in revised form 9 April 2017; Accepted 3 May 2017
Available online 05 May 2017

2. Materials and methods

2.1. Human subjects

Patients were evaluated by a board certified child neuropsychiatrist (SKA) and underwent brain MRI to assess any abnormal anatomical features. The Arab family was recruited from the pediatric clinic in King Abdullah University hospital in the northern Jordan. The parents were first cousins and had two affected female siblings and five unaffected female siblings. Ethnically matched normal controls (257 individuals) were tested for testing the presence and frequency of the mutation in the local population. Informed consent was obtained from all participating individuals or their legal guardians. The study was approved by the Institutional review board in both the Jordan University of Science and Technology and the Qatar Biomedical Research Institute at Hamad bin Khalifa University. Peripheral blood or saliva was collected from all participants for DNA extraction by Gentra Puregene Blood Kit (Cat# 158422, QIAGEN, Germantown, MD, USA) for blood and by Oragene collection kits (OGR-500, Oragene, Ottawa, Canada) for saliva as per the manufacturer's instructions.

2.2. Autozygosity mapping and exome sequencing

Autozygosity mapping and whole exome sequencing was carried out as described previously (Alkhateeb et al., 2016). Briefly, stretches of autozygosity > 2 Mb were used as surrogates of autozygosity in the affected individual; these stretches were heterozygous in parents and not homozygous mutant in unaffected siblings. This was determined using Illumina human mapping HumanOmniExpress-12 v1.0 array (Illumina, Inc., San Diego, CA, USA) following the manufacturer's instructions. Copy number variants (CNVs) and genotypes were analyzed using GenomeStudio software v2011.1 (Illumina, San Diego, CA, USA). Autozygosity intervals were determined by HomozygosityMapper (Seelow and Schuelke, 2012). Candidate genes within the autozygosity intervals (*GFRA3*, *GRIN3B* and *PTBP1*) were sequenced by standard Sanger sequencing for all gene exons on ABI 3730 automatic DNA sequencer (Applied Biosystems, Foster city, CA, USA). Primers used to assay for the *PTBP1* variant where (forward) 5′-TACCCTGTGACCCTGGATGT-3′ and (reverse) 5′-GACTGAGCGGCCCTAATAGA-3′, based on ref. seq NM_002819. Details of other primers used to screen the full *PTBP1* gene in addition to *GFRA3*, and *GRIN3B* genes are available upon request. Whole exome sequencing was then performed for all participating family members. Enrichment was performed using the Nextera Rapid Capture exome kit V1.2 and run on an Illumina HiSeq 2500 (Illumina, Inc., San Diego, CA, USA). Mean target coverage for all samples was 108× with 87% of bases covered at > 20×. Detailed coverage of each individual is provided in the Supplementary material. Mapping and analysis of variants was done as described before (Alkhateeb et al., 2016). Validation of whole exome identified variants, confirming of variant co-segregation, and testing of local population variant frequency in 257 controls was performed by standard Sanger sequencing. Candidate variant tested after whole exome sequencing (WES) was *CLMN* p.k863*. Primers used to screen these variants are listed in the Supplementary material. Primers used to assay for *CLMN* variant; forward: 5′-GCAAACCCCCTAGAAGAAAA-3′, reverse: 5′-TGAACACTTCCTGGTGCTACA-3′ (ref. seq: NM_024734).

2.3. In silico analysis

Polyphen-2, SIFT, and MutationTaster were used to predict the effect of genetic variants. The web-based ESEfinder 3.0 (http://rulai.cshl.edu/tools/ESE) was used to assay the effect of *PTBP1* variant on the binding of serine/arginine-rich (SR) proteins SRSF1, SRSF1 (IgM-BRCA1), SRSF2, SRSF5, and SRSF6 (Cartegni et al., 2003) to the DNA segment containing the variant. These proteins are members of the splicing factors family that play key roles in splicing. This tool facilitates analysis of exonic sequences to predict if nucleotide variations can disrupt binding of SR proteins and thus points to their effect as probable exonic splicing enhancers/repressors. Additionally, Human splicing finder, HSF 3.0 (http://www.umd.be/HSF3/) was also used to assess the variant effect (Desmet et al., 2009).

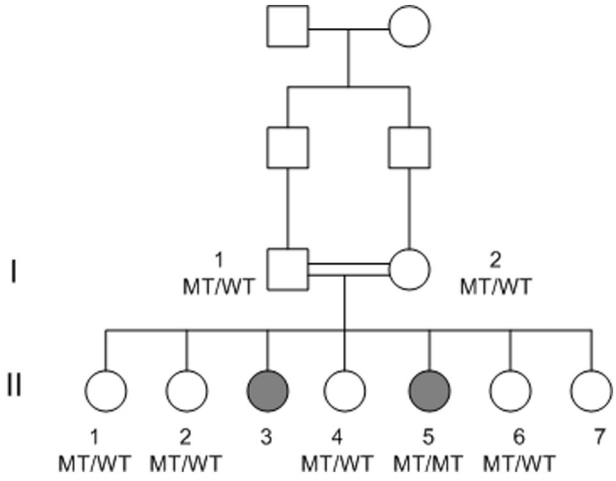

Fig. 1. Pedigree of the family. Shaded symbols denote affecteds. Double lines denote consanguineous marriage. Circles are for females and squares are for males. Generations are denoted by roman numerals and individuals by Arabic numerals. MT stands for mutant allele (*PTBP1*, c.708C > T) and WT for wild type allele.

3. Results

Here we have identified a first-cousin consanguineous family of Arab descent. The family had two affected daughters and 5 unaffected daughters (Fig. 1). The patient presented initially with mild developmental delay, however with apparently progressive course. She was ultimately ambulatory, however with failure to gain cognitive and language skills, and development of intellectual disability. She also developed progressive myoclonic epilepsy, and electroencephalography was consistent with Lennox Gastuat syndrome. MRI showed transient subtle globus pallidi signal abnormalities that subsided with repeat MRI, last MRI was normal. Opthalmological evaluation was normal, basic metabolic screen was normal. Her sister showed a similar clinical of global developmental delay; she was ambulatory, but also had progressive intellectual disability and epilepsy. Patients had a paternal cousin with the same clinical picture. Samples were collected from one of the patients, her parents and her 4 unaffected sisters. All samples underwent genome-wide SNP genotyping. Analysis of CNVs did not detect potential pathogenic co-segregating variants. Autozygosity mapping was done for the family, 10 homozygous intervals were detected in the patient, heterozygous in parents and non-homozygous mutant in the four unaffected siblings (Fig. 2, Table 1). Although the intervals spanned 140 Mb we elected to screen 3 genes for their probable functional candidacy, *GFRA3*, *GRIN3B*, and *PTBP1* genes. One novel synonymous variant stood out in *PTBP1* gene, c.708C > T (p.H236H, rs762173288) (Fig. 3). This nucleotide change was extremely rare with only one allele found out of 120,782 alleles in the Exome Aggregation Consortium (ExAC) (URL: http://exac.broadinstitute.org) (accessed May 2016).

To gain a more comprehensive data for other variants that might be responsible for the phenotype, we subjected 6 samples for WES (proband, parents, and 3 unaffected sibs). Overall, mean target coverage was 108× with a mean 87% covered > 20× (for detailed coverage see methodology). A total of 152,821 variants were generated, after subjection to filtration, 5 variants remained (see Supplementary table). Of those 5 variants, one was selected for follow up (p.K863* in *CLMN* gene), the other genes were excluded as they were associated with other phenotypes, have a function irrelevant to the phenotype, and/or were