European Journal of Human Genetics (2011) 19, 717–720

www.nature.com/ejhg

SHORT REPORT

Hybridisation-based resequencing of 17 X-linked intellectual disability genes in 135 patients reveals novel mutations in *ATRX*, *SLC6A8* and *PQBP1*

Lars R Jensen[1,13], Wei Chen[1,13], Bettina Moser[1], Bettina Lipkowitz[1], Christopher Schroeder[2], Luciana Musante[1], Andreas Tzschach[1], Vera M Kalscheuer[1], Ilaria Meloni[3], Martine Raynaud[4], Hilde van Esch[5], Jamel Chelly[6], Arjan PM de Brouwer[7], Anna Hackett[8], Sigrun van der Haar[9], Wolfram Henn[10], Jozef Gecz[11,12], Olaf Riess[2], Michael Bonin[2], Richard Reinhardt[1], Hans-Hilger Ropers[1] and Andreas W Kuss*[,1]

X-linked intellectual disability (XLID), also known as X-linked mental retardation, is a highly genetically heterogeneous condition for which mutations in >90 different genes have been identified. In this study, we used a custom-made sequencing array based on the Affymetrix 50k platform for mutation screening in 17 known XLID genes in patients from 135 families and found eight single-nucleotide changes that were absent in controls. For four mutations affecting *ATRX* (p.1761M>T), *PQBP1* (p.155R>X) and *SLC6A8* (p.390P>L and p.477S>L), we provide evidence for a functional involvement of these changes in the aetiology of intellectual disability.
European Journal of Human Genetics (2011) **19,** 717–720; doi:10.1038/ejhg.2010.244; published online 26 January 2011

Keywords: X-linked intellectual disability; X-linked mental retardation; array-based resequencing; mutation analysis; automated PCR

INTRODUCTION

Mental retardation/intellectual disability (ID) has a prevalence of ~2% in the general population.[1] Therefore, as well as because of the life-long restrictions for the patients and their families in combination with extremely limited therapeutic options, ID ranges among the top health issues of our time. Significantly more males compared with females are affected, which can be partly accounted for by mutations in genes on the X chromosome. Moderate-to-severe forms of X-linked ID (XLID) affect 0.05%[2] of the population in developed countries. The majority of the patients show non-syndromic forms of XLID (NS-XLID), in which ID is the only clinically consistent manifestation.

On the genetic level, NS-XLID is highly heterogeneous and the causative mutations known to date were found in more than 37 genes, each gene accounting for only a few percent of the cases.[1,3] Families with several affected members from different generations are suitable for linkage analysis, and such families have therefore been instrumental in finding many of the presently known mutations that cause NS-XLID. However, to identify the underlying genetic defect in small ID families or in sporadic patients, a different approach is required. Thus, to investigate the genetic basis of ID in 135 small families from the cohort of the European MRX (Euro-MRX) consortium (http://www.euromrx.com), we have designed a resequencing array based on the Affymetrix 50k platform, containing the coding and splice site regions of 17 XLID genes (see below), with a relatively high mutation frequency in NS-XLID patients. Application of this array led to the identification of eight previously unknown changes in these genes.

SUBJECTS AND METHODS

Patients and controls

DNA from 135 unrelated ID patients without a molecular diagnosis were collected through the Euro-MRX consortium. The majority of patients were from families with no more than two affected brothers. After obtaining written informed consent from the probands or their parents, DNA was extracted from patient blood or blood-derived lymphoblastoid cell lines using standard methods. The study was approved by the relevant institutional review board.

Resequencing array

A custom resequencing array was designed based on the Affymetrix 50k platform (Affymetrix, Santa Clara, CA, USA), allowing up to 50 kb of DNA to be sequenced in both orientations in a single hybridisation experiment.

[1]Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany; [2]Department of Medical Genetics, University of Tübingen, Tübingen, Germany; [3]Medical Genetics Unit, Department of Molecular Biology, University of Siena, Siena, Italy; [4]Services de Génétique (INSERM), Tours, France; [5]Centre for Human Genetics, University Hospital Leuven, Leuven, Belgium; [6]Institut Cochin de Génétique Moléculaire, Centre National de la Recherche Scientifique (INSERM), Paris, France; [7]Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences and Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; [8]The GOLD Service, Hunter Genetics, Waratah, New South Wales, Australia; [9]Practice for Gynaecology and Medical Genetics, Nuremberg, Germany; [10]Institute of Human Genetics, University Hospital Homburg/Saarland, Homburg/Saar, Germany; [11]Department of Genetics and Molecular Pathology, SA Pathology at the Women's and Children's Hospital, University of Adelaide, Adelaide, Australia; [12]Department of Paediatrics, University of Adelaide, Adelaide, Australia
*Correspondence: Dr AW Kuss, Institute for Human Genetics, Interfaculty Institute for Genetics and Functional Genomics, Ernst Moritz Arndt University of Greifswald, Fleischmannstr. 42-44, 17475 Greifswald, Germany.
Tel: +49 (0)3834 86 5371; Fax: +49 (0)3834 86 5393; E-mail: andreas.kuss@uni-greifswald.de
[13]These authors contributed equally to this work.
Received 29 April 2010; revised 1 December 2010; accepted 2 December 2010; published online 26 January 2011

The oligonucleotides on the array contain sequences that include the coding regions and splice sites (10 nucleotides flanking each exon) of *ACSL4*, *ARX*, *ATRX*, *DLG3*, *FTSJ1*, *GDI1*, *IL1RAPL1*, *JARID1C*, *MECP2*, *NLGN4*, *PAK3*, *PHF6*, *PHF8*, *PQBP1*, *SLC6A8*, *TM4SF2* and *ZNF41*.

PCR amplification of DNA, purification and hybridisation

A total of 151 PCR amplicons were produced from genomic DNA (Supplementary Material) and subsequently purified by alcohol. The purified DNA fragments were pooled in equimolar quantities (0.055 pmol of each PCR product, that is, 36.3 ng per 1000 bp) and the volume was reduced by evaporation using a vacuum centrifuge.

Fragmentation of the products, labelling of the products with biotin, hybridisation, washing and scanning procedures were carried out according to the CustomSeq Resequencing Array Protocol version 2.0 (Affymetrix).

Data analysis

The raw sequencing data were analysed using the Affymetrix GeneChip Analysis Software (GSEQ) version 4.0 (Affymetrix). Only nucleotides that were called on both DNA strands using a quality score threshold equal to 0 were considered. As we only used DNA from male patients (hemizygous for X-chromosomal genes), we did not consider heterozygous base calls. The entire data set was also analysed using SeqC (JSI medical systems, Kippenheim, Germany) as previously described.[4]

RT-PCR and measurement of ACSL4 and SLC6A8 activity

RT-PCR was performed according to standard protocols. Enzymatic activity of ACSL4 and the urinary creatine/creatinine ratio were measured as described previously.[5,6]

Protein extraction and western blot analysis

Cells from control and patient lymphoblastoid cell lines were lysed in buffer containing 48% urea, 15 mM Tris (pH 7.5), 8.7% glycerol, 1% SDS, 0.004% bromophenol blue and 143 mM β-mercaptoethanol. Western blot analysis was performed as previously described.[7]

RESULTS AND DISCUSSION

Using a combination of automated and manual PCR, we produced 151 amplicons, containing the coding regions and splice sites of 17 known XLID genes (Table 1), from each of 135 ID patient DNAs (Supplementary Table 1).

As all PCR products had the expected size, we could exclude the presence of larger deletions or insertions in these 17 genes. Small in-dels cannot be detected because of the nature of hybridisation-based sequencing,[8,9] however, as in-dels in the 17 investigated genes are much less frequent than missense mutations (Human Mutation Database), it is possible, but not very likely, that undetected in-dels are present in these 17 genes in the analysed patient cohort.

In addition, we sequenced the amplicons containing the tri-nucleotide repeats in *ARX* by Sanger sequencing to rule out mutations in this region, as repeat expansions are also not detectable by array-based resequencing.

The remaining PCR products were submitted to fragmentation and labelling. After hybridisation, washing and scanning of the arrays, we obtained call rates between 90 and 96% and found 478 different nucleotide changes (excluding known SNPs). The specified call accuracy for this array type is 99.99% (1 error in 10 000 nucleotide calls) when – like in this study – no heterozygous base calls are expected. Thus, four to five falsely called nucleotides per patient DNA (600 in 135 patients) had to be expected. As a filtering criterion to remove false-positive sequence changes from the data set, we used the call rate among the five immediately adjacent bases on either side of a given base. Assuming that fragments with poor sequence quality harboured a substantial amount of falsely called bases, we focused on sequence changes for which the 10 surrounding nucleotides contained at least 6 positions with a base call. This threshold reduced the number of putative nucleotide changes from 478 to 75. The amplicons containing these changes were then investigated by Sanger sequencing, which led to the identification of 10 verifiable nucleotide substitutions in 9 families. All these sequence changes were found in fragments in which at least 8 called bases were present in the 10 surrounding nucleotide positions.

To test the validity of the filtering procedure, we applied SeqC, a module of the *Sequence Pilot* software (JSI medical systems), for an independent analysis of the whole data set, and identified one

Table 1 Genes represented on the custom-made resequencing array

Symbol	*Reference sequence*	*OMIM no.*	*Gene name*	*Chromosomal location*	*Amplicons*[a]
ARX	NM_139058.1	300382	Aristaless-related homeobox	Xp21.3	6
JARID1C	NM_004187.1	314690	Jumonji, AT-rich interactive domain 1C	Xp11.22–p11.21	13
SLC6A8	NM_005629.1	300036	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8	Xq28	2
DLG3	NM_021120.1	300189	Discs, large homolog 3 (neuroendocrine-dlg, Drosophila)	Xq13.1	13
GDI1	NM_001493.1	300104	GDP dissociation inhibitor 1	Xq28	6
PAK3	NM_002578.2	300142	p21 (CDKN1A)-activated kinase 3	Xq22.3	11
IL1RAPL1	NM_014271.2	300206	Interleukin 1 receptor accessory protein-like 1	Xp22.1–p21.3	10
MECP2	NM_004992.2	300005	Methyl CpG-binding protein 2 (Rett syndrome)	Xq28	3
TSPAN7	NM_004615.2	300096	Tetraspanin 7	Xq11	7
ACSL4	NM_022977.2	300157	Acyl-CoA synthetase long-chain family member 4	Xq22.3–q23	7
PQBP1	NM_005710.2	300463	Polyglutamine-binding protein 1	Xp11.23	3
ZNF41	NM_007130.1	314995	Zinc finger protein 41	Xp11.23	4
NLGN4	NM_020742.2	300427	Neuroligin 4, X-linked	Xp22.33	5
FTSJ1	NM_177439.1	300499	FtsJ homolog 1 (*E. coli*)	Xp11.23	3
ATRX	NM_000489.3	300032	Alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, *S. cerevisiae*)	Xq21.1	31
PHF8	NM_015107.1	300560	PHD finger protein 8	Xp11.22	16
PHF6	NM_032458.2	300414	PHD finger protein 6	Xq26	6

[a]Number of amplicons required to cover the coding region and splice sites of the respective genes.

additional change in the intron of *ACSL4*. This change was not detected with the previously applied filtering procedure, because it is located close to the end of an amplicon and, therefore, lacked the mandatory five nucleotides on one side.

For each verified nucleotide alteration, we then sequenced DNA from all available family members to investigate co-segregation with the disorder. Furthermore, we analysed DNA from 138 healthy males to exclude common polymorphisms. This ruled out the change in *PHF8* c.441T>G, which was found once in the control panel, and the change in *NLGN4X* c.968A>G, which did not segregate with the disorder in the family. These two changes are therefore unlikely to be pathologically relevant. The remaining eight potentially disease-causing changes are listed in Table 2.

For the silent change in *JARID1C* and the intronic change in *ACSL4*, RT-PCR experiments showed normal levels of gene expression and a normal splicing pattern in patient lymphoblastoid cell lines (data not shown), indicating that these changes are probably benign. What is more, in case of the missense change in *ACSL4* (c.1382C>T, p.461P>L), the functional integrity of the gene product seemed uncompromised, as enzymatic activity of ACSL4 in a lymphoblastoid cell line from affected individual was normal. In addition to the intronic change in ACSL4, we also found an intronic change in *PQBP1* in the same patient DNA (P122). The intronic change (c.586+25G>A) in the main known *PQBP1* transcript (NM_005710.2) did not alter the *PQBP1* splicing pattern, as investigated by RT-PCR. However, the change also introduces a missense change (p.201G>D) in one *PQBP1* isoform (AJ973600.1). This isoform is very weakly expressed, and further studies are required to determine whether this change is pathogenic or not.

The patient in whom we found a p.1761M>T missense change in *ATRX* was previously considered to have a non-syndromic phenotype. As our finding indicated that he was actually suffering from ATRX (Alpha-thalassemia X-linked mental retardation) syndrome, we initiated a clinical re-examination. This revealed the presence of characteristic dysmorphic features, providing strong evidence for the pathological relevance of this novel mutation.

One change in *SLC6A8* (c.1169C>T, p.390P>L) is identical to a change previously reported in a family with ID[10] and later shown to be pathogenic.[11]

The second change (c.1430C>T, p.477S>L) in *SLC6A8* was not observed before, and investigation of the urine creatine/creatinine ratio in the index patient revealed an abnormally elevated value of 1.01 (normal levels in adults: <0.25), confirming the molecular finding.

Interestingly, we also identified a novel nonsense mutation (c.463C>T) in exon 4 of *PQBP1*, which is frequently mutated in ID patients.[12] The change results in the introduction of a premature stop codon (p.155R>X). Western blot analysis demonstrated that truncated PQBP1 protein resulting from this mutation is present in the patient (Figure 1), as already shown for other patients with similar *PQBP1* mutations,[7] supporting its pathological relevance.

In summary, the discovery of 4 putative disease-causing mutations in 7 out of 135 XLID patients equals a proportion of ~3%, which is well within the expected range. First because, on the basis of observations by the EURO-MRX consortium, only ~40% of the male patients (ie, here 54 out of 135) from small families can be assumed to carry a single gene defect on the X chromosome[13] and second because the frequencies of disease-causing mutations in the selected genes are each below a few percent.[1,14] The fact that we found an excess of missense mutations, as compared with protein-truncating mutations, is also in

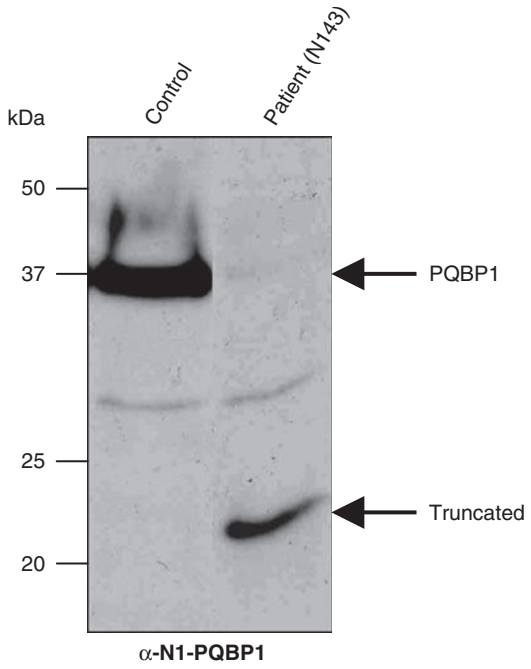

Figure 1 PQBP1 expression in control and patient lymphoblastoid cell lines. Whole-cell lysates from a control and a patient lymphoblastoid cell line that harbours a truncating mutation in *PQBP1* (N143) were run in parallel on an SDS-PAGE gel. The gel was blotted and probed with an antibody specific for the N-terminal part of PQBP1 (α-N1-PQBP1).[7] In the control cell line lysate, a band corresponding to wild-type PQBP1 protein was observed at ~37 kDa, and in the patient cell line lysate, a truncated PQBP1 protein was observed at ~22 kDa.

Table 2 DNA changes found in MR patients, but not in dbSNP or in controls (*n*=138)

Gene	*Reference sequence*	*Nucleotide change*	*Amino acid change*	*Family*	*Co-segregation*	*Pathological relevance*
ACSL4	NM_022977.2	c.1382C>T	p.461P>L	A100	Co-segregating	Benign
ACSL4	NM_022977.2	c.1855+11A>G		P122	No DNA available	Benign
ATRX	NM_000489.3	c.5282T>C	p.1761M>T	D020	Co-segregating	Causative
JARID1C	NM_004187.2	c.1827C>T	p.609Y>Y	T056	No DNA available	Benign
PQBP1	NM_005710.2	c.463C>T	p.155R>X	N143	Co-segregating	Causative
PQBP1	NM_005710.2	c.586+25G>A		P122	No DNA available	Probably benign
SLC6A8	NM_005629.3	c.1169C>T	p.390P>L	D077	Co-segregating	Causative[a]
SLC6A8	NM_005629.3	c.1430C>T	p.477S>L	L088	Co-segregating	Causative

[a]The change c.1169C>T in SLC6A8 was reported by Rosenberg *et al*.[10]

line with previous findings.[14] The advantages of array-based resequencing include the potential for automation of sample preparation, the speed with which base calling is performed and the comparatively low cost. Array-based resequencing is therefore an option when large cohorts are analysed for sequence changes in defined DNA regions. Considerations for array design and the limitations of array-based resequencing have been thoroughly discussed.[15] However, alternatives to array-based resequencing are now available in the form of next-generation sequencing methods, which use sequencing by synthesis. Recently, a protocol for enrichment and subsequent sequencing of DNA fragments that contain the desired sequences has been published.[16] This or similar approaches can reduce the time involved in the sequencing procedure considerably and will provide new options for elucidating the aetiology of ID in the so far unresolved cases.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Sven Poths for technical assistance. This study was supported by the Max Planck Innovation Fund. Additional funding was provided to HHR by the German Federal Ministry of Education and Research (MRNET 01GS08161-2).

DATABASES

UCSC Genome Browser: http://genome.ucsc.edu/cgi-bin/hgTracks?org=human
The Human Gene Mutation Database: http://www.hgmd.cf.ac.uk/ac/validate.php
Ensembl: http://www.ensembl.org/index.html
Euro-MRX Consortium: http://www.euromrx.com

1 Ropers HH: Genetics of intellectual disability. *Curr Opin Genet Dev* 2008; **18**: 241–250.
2 Leonard H, Wen X: The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* 2002; **8**: 117–134.
3 Gécz J, Shoubridge C, Corbett M: The genetic landscape of intellectual disability arising from chromosome X. *Trends Genet* 2009; **25**: 308–316.
4 Schroeder C, Stutzmann F, Weber BH, Riess O, Bonin M: High-throughput resequencing in the diagnosis of BRCA 1/2 mutations using oligonucleotide resequencing microarrays. *Breast Cancer Res Treat* 2010; **122**: 287–297.
5 Longo I, Frints SG, Fryns JP *et al*: A third MRX family (MRX68) is the result of mutation in the long chain fatty acid-CoA ligase 4 (FACL4) gene: proposal of a rapid enzymatic assay for screening mentally retarded patients. *J Med Genet* 2003; **40**: 11–17.
6 Almeida LS, Verhoeven NM, Roos B *et al*: Creatine and guanidinoacetate: diagnostic markers for inborn errors in creatine biosynthesis and transport. *Mol Genet Metab* 2004; **82**: 214–219.
7 Musante L, Kunde SA, Sulistio TO *et al*: Common pathological mutations in PQBP1 induce nonsense-mediated mRNA decay and enhance exclusion of the mutant exon. *Hum Mutat* 2010; **31**: 90–98.
8 Hartmann A, Thieme M, Nanduri LK *et al*: Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. *Hum Mutat* 2009; **30**: 115–122.
9 Fokstuen S, Lyle R, Munoz A *et al*: A DNA resequencing array for pathogenic mutation detection in hypertrophic cardiomyopathy. *Hum Mutat* 2008; **29**: 879–885.
10 Rosenberg EH, Almeida LS, Kleefstra T *et al*: High prevalence of SLC6A8 deficiency in X-linked mental retardation. *Am J Hum Genet* 2004; **75**: 97–105.
11 Rosenberg EH, Martínez Muñoz C, Betsalel OT *et al*: Functional characterization of missense variants in the creatine transporter gene (SLC6A8): improved diagnostic application. *Hum Mutat* 2007; **28**: 890–896.
12 Kalscheuer VM, Freude K, Musante L *et al*: Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nat Genet* 2003; **35**: 313–315.
13 de Brouwer AP, Yntema HG, Kleefstra T *et al*: Mutation frequencies of X-linked mental retardation genes in families from the EuroMRX consortium. *Hum Mutat* 2007; **28**: 207–208.
14 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.
15 Kothiyal P, Cox S, Ebert J, Aronow BJ, Greinwald JH, Rehm HL: An overview of custom array sequencing. *Curr Protoc Hum Genet* 2009 Chapter 7:Unit 7.17.
16 Hodges E, Rooks M, Xuan Z *et al*: Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nat Protoc* 2009; **4**: 960–974.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)