European Journal of Human Genetics (2007) 15, 68–75

www.nature.com/ejhg

ARTICLE

X-linked mental retardation: a comprehensive molecular screen of 47 candidate genes from a 7.4 Mb interval in Xp11

Lars Riff Jensen[1,7], Steffen Lenzner[1,7], Bettina Moser[1], Kristine Freude[1], Andreas Tzschach[1], Chen Wei[1], Jean-Pierre Fryns[2], Jamel Chelly[3], Gillian Turner[4], Claude Moraine[5], Ben Hamel[6], Hans-Hilger Ropers*[,1] and Andreas Walter Kuss[1]

[1]*Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany;* [2]*Centre for Human Genetics, University Hospital Leuven, Leuven, Belgium;* [3]*Département de Génétique et Pathologie Moléculaire, Institut Cochin, Paris, France;* [4]*Hunter Genetics and University of Newcastle, New South Wales, Australia;* [5]*Services de Génétique-INSERM U316, CHU Bretonneau, Tours, France;* [6]*Department of Human Genetics, University Medical Centre, Nijmegen, The Netherlands*

About 30% of the mutations causing nonsyndromic X-linked mental retardation (MRX) are thought to be located in Xp11 and in the pericentromeric region, with a particular clustering of gene defects in a 7.4 Mb interval flanked by the genes *ELK1* and *ALAS2*. To search for these mutations, 47 brain-expressed candidate genes located in this interval have been screened for mutations in up to 22 mental retardation (MR) families linked to this region. In total, we have identified 57 sequence variants in exons and splice sites of 27 genes. Based on these data, four novel MR genes were identified, but most of the sequence variants observed during this study have not yet been described. The purpose of this article is to present a comprehensive overview of this work and its outcome. It describes all sequence variants detected in 548 exons and their flanking sequences, including disease-causing mutations as well as possibly relevant polymorphic and silent sequence changes. We show that many of the studied genes are unlikely to play a major role in MRX. This information will help to avoid duplication of efforts in the ongoing endeavor to unravel the molecular causes of MRX.

European Journal of Human Genetics (2007) **15,** 68–75. doi:10.1038/sj.ejhg.5201714; published online 13 September 2006

Keywords: MRX; mutation screening; X-linked mental retardation; Xp11

Introduction

Mental retardation (MR), defined by an intelligence quotient below 70, is characterized by a global deficiency in cognitive abilities. It represents the most frequent phenotypic manifestation of abnormal development in the central nervous system, affecting about 2% of the population in industrialized countries.[1] X-linked MR affects approximately 1 out of 1000 males, $\frac{2}{3}$ of which show nonsyndromic X-linked MR (MRX). MRX is a genetically and clinically heterogeneous disorder with MR as the only clinically consistent feature. To date >20 genes playing a role in MRX have been identified.[2] However, the mutation frequency in each gene is low, which implies that the majority of genetic defects involved in MRX remains to be identified.

Analysis of published as well as unpublished linkage data from numerous families has shown that mutations

*Correspondence: Professor H-H Ropers, Max Planck Institute of Molecular Genetics, Ihnestrasse 73, Berlin 14195, Germany.
Tel: +49 3084131241; Fax: +49 3084131383;
E-mail: ropers@molgen.mpg.de
[7]These authors contributed equally to this work
Received 26 April 2006; revised 26 July 2006; accepted 27 July 2006; published online 13 September 2006

involved in MRX seem to cluster in three different regions of the human X-chromosome.[3] Two regions, one at Xq28 and the other at Xp22.1–p21.3, contain known MR genes. The presence of a third broad peak on proximal Xp suggested the existence of novel MR genes in this region. Therefore, we have chosen a 7.4 Mb interval in Xp11, flanked by the genes *ELK1* and *ALAS2*, to systematically screen the coding regions and splice sites of 47 candidate genes in between for mutations in up to 22 unrelated XLMR families with overlapping linkage intervals. Based on this study, four MR genes have been identified and published, but many other sequence variants, with so far unknown functional relevance have been identified as well.

The purpose of this article is to provide a comprehensive overview about the outcome of this work. It describes all sequence variants in a total of 548 exons and their flanking sequences.

Materials/subjects and methods

Patients and controls

Families with linkage intervals overlapping the region between *ELK1* and *ALAS2* on proximal Xp were selected. These included patients from five different families with syndromic forms of X-linked MR, one patient with Sutherland–Haan syndrome,[4] and one patient with Wieacker–Wolf syndrome.[5] Furthermore, 20 unrelated MRX patients were chosen, including eight patients from previously described MRX families: MRX15,[6] MRX18,[7] MRX26,[8] MRX44, MRX45, MRX52,[9] MRX55[10] and MRX65.[11] A survey of the investigated families, their individual linkage intervals as well as additional clinical features is given in Table 1. A different set consisting of 180 individuals from small families with presumed MRX was obtained through the Euro-MRX Consortium. These families show two to five affected brothers but no obligate female carriers. As controls, 168 unrelated male individuals, either students or healthy blood donors, have been used. For *FLJ14103*, *HDAC6*, *SLC38A5*, *TRO* and *ZNF41*, additional controls including also females were used, and for *PLP2*, DNA samples from patients with other disorders but without MR were included as well.

DNA extraction

Blood samples and lymphoblastoid cell lines (LCL) from patients and controls were obtained through the Euro-MRX Consortium (http://www.euromrx.com/). Genomic DNA was extracted using standard methods.

Table 1 Linkage intervals and additional features of 27 XLMR families

EURO-MRX family ID	*Linkage intervals*[a]	*Cytogenetic bands*	*Additional clinical features*	*Euro-MRX number*	*References*
A005	**DXS1003–DXS990**	**Xp11.3–Xq21.32**			
A006	**DXS1214–DXS1126**	**Xp21.2–Xp11.23**		**MRX18**	**Gedeon *et al*[7]**
A007	**DXS1003–DXS990**	**Xp11.3–Xq21.32**			
A008	**DXS1202–DXS990**	**Xp21.3–Xq21.32**			
P001	**DXS1068–DXS1275**	**Xp11.4–Xq13.1**		**MRX55**	**Deqaqi *et al*[10]**
D002	DXS1068–DXS991	Xp11.4–Xp11.21			
D043	DXSMAOB–DXS54	Xp11.3–Xq21.2		MRX26	Robledo *et al*[8]
L038	DXS1003–DXS8020	Xp11.3–Xq22.1			
N001	DXS1055–ALAS2	Xp11.3–Xp11.21		MRX44	Hamel *et al*[9]
N002	DXS1003–ALAS2	Xp11.3–Xp11.21		MRX45	Hamel *et al*[9]
N010	DXS573–DXS990	Xp11.23–Xq21.32		MRX65	Yntema *et al*[11]
N017	DXS1003–PGKP1	Xp11.3–Xq12			
N018	ALAS2–DXS3	Xp11.21–Xq21.32		MRX52	Hamel *et al*[9]
N045	DXS7–DXS453	Xp11.3–Xq13.1			
T003	MAOB–DXS3	Xp11.3–Xq21.32			Raynaud *et al*[12]
T008	DXS164–DXS988	Xp21.1–Xp11.22		MRX15	Raynaud *et al*[6]
T025	DXS1214–DXS1212	Xp21.2–Xq25			
T040	DXS556–DXS1001	Xp11.4–Xq24			Raynaud *et al*[12]
T050	DXS426–DXS1106	Xp11.23–Xp22.2			Raynaud *et al*[12]
T102	DXS997–ALAS2	Xp21.1–Xp11.21			
L017	DXS1214–DXS990	Xp21.2–Xq21.32	Spastic paraplegia		Claes *et al*[13]
L045	DXS1214–DXS991	Xp21.2–Xp11.21	Spastic paraplegia		
N009	DXS989–HPRT	Xp22.1–Xq26.2	Microcephaly, growth retardation, cleft uvula		
N040	DXS426–DXS3	Xp11.23–Xq21.32	Heart defect, cleft palate, short stature		Hamel *et al*[9]
N042	DXS337–DXS990	Xp11.3–Xq21.32	Cleft lip/palate		Siderius *et al*[14]
[SHS]	MAOA–DXS1125	Xp11.3–Xq13.1	Sutherland–Haan Syndrome		Sutherland *et al*[4]
[WWS]	PFC–DXS339	Xp11.23–Xq13.1	Wieacker–Wolff Syndrome		Kloos *et al*[5]

Bold: Families investigated only by Northern blot analysis.
[a]Flanking recombinant markers.

Selection of cDNA sequences for molecular screen

At the start of the project, 65 RefSeq genes and many more mRNA and EST sequences encoding hypothetical proteins (UCSC Genome Browser; April 2002 Freeze) were known in the candidate region at Xp11.23–11.21 flanked by the genes *ELK1* and *ALAS2*. Candidate genes (Table 2) were selected using the following criteria: firstly, gene expression in nervous tissue (indicated by ESTs present in the Unigene database), secondly, no involvement in other disorders without MR and thirdly, presence of an open-reading frame. The 47 genes fulfilling these criteria are listed in Table 2.

Primer design

Primers for amplification of the entire coding regions and exon–intron boundaries of all genes listed in Table 2 were

Table 2 Alphabetical list of 47 screened candidate genes from the Xp11 interval *ELK1-ALAS2*

Gene		*mRNA accession no. (sequence used for primer design)*	*Total number of patients (max: 22)*
Abbreviation	*Full name*		
ALAS2	Aminolevulinate, delta-, synthase 2	NM_000032	16[a]
APEX2	**Apurinic/apyrimidinic endonuclease 2**	**AJ011311**	**16[a]**
CCNB3	Cyclin B3 isoform 3	NM_033031	18[b]
DT1P1A10	Hypothetical protein DT1P1A10	BC011825	15[a]
ELK1	**ELK1 protein/member of ETS oncogene family**	**NM_005229**	**17[a,b]**
FGD1	Faciogenital dysplasia protein (Aarskog–Scott syndrome)	NM_004463	15[a]
FLJ10613	Hypothetical protein FLJ10613	BC011720	15[a]
FLJ14103	Hypothetical protein FLJ14103	NM_024689	21[a]
FLJ20344	**Hypothetical protein FLJ20344**	**NM_017776**	**21[a]**
FLJ21687	Hypothetical protein FLJ21687	AK025340	18[b]
FTSJ1	**FtsJ homolog 1 isoform a**	**NM_012280**	**22**
GRIPAP1	**GRIP1-associated protein 1**	**AB032993**	**18[b]**
GSPT2	G1 to S phase transition 2/peptide chain release factor 3	NM_018094	18[b]
HADH2	**Hydroxyacyl-Coenzyme A dehydrogenase, type II**	**AF035555**	**20[a]**
HDAC6	**Histone deacetylase 6**	**NM_006044**	**16[a]**
JARID1C	**Smcx homolog, X-linked (mouse)**	**L25270**	**22**
JM1	**JM1 protein**	**BC000972**	**18[b]**
JM11	**Hypothetical protein FLJ31728/JM11 protein**	**NM_033626**	**16[a]**
JM4	**JM4 protein**	**BC021213**	**12[a,b]**
JM5	**JM5 protein**	**AJ005897**	**12[a,b]**
KIAA0522	Hypothetical protein KIAA0522	AB011094	16[a,b]
KIAA1202	Hypothetical protein KIAA1202	AB033028	17[a]
LMO6	**LIM domain only 6**	**NM_006150**	**15[a]**
MAGED2	**Melanoma antigen, family D, 2**	**AF128528**	**21[a]**
MAGE-E1	Hypothetical protein (MAGE-E1a protein)	NM_030801	18[b]
MG61/PPN	Porcupine	NM_022825	18[b]
PCSK1N[c]	Proprotein convertase subtilisin/kexin type 1	NM_013271	11[a,b]
PHF8	**Hypothetical protein KIAA1111**	**AB029034**	**18[b]**
PIM2	**Pim-2 oncogene**	**U77735**	**18[b]**
PLP2	**Proteolipid protein 2 (colonic epithelium-enriched)**	**L09604**	**15[a,b]**
PPP1R3F	Protein phosphatase 1, regulatory (inhibitor) subunit 3F	NM_033215	18[b]
PQBP1	**Polyglutamine binding protein 1**	**AJ242829**	**22**
PRKWNK3	Protein kinase WNK3	AJ409088	13[a]
PRO0659	Hypothetical protein PRO0659	NM_014138	21[a]
RBM3	RNA binding motif protein 3	NM_006743	16[a,b]
SLC35A2	Solute carrier family 35 (UDP-galactose transporter), member A2	D88146	15[a]
SLC38A5	Amino acid transport system N2	AF276889	12[a,b]
SMC1L1	SMC1 structural maintenance of chromosomes 1-like 1 (yeast)	NM_006306	18[b]
SUV39H1	**Suppressor of variegation 3–9 homolog 1 (*Drosophila*)**	**NM_003173**	**18[b]**
T54	**T54 protein**	**NM_015698**	**16[a,b]**
TFE3	**Transcription factor binding to IGHM enhancer 3**	**NM_006521**	**18[b]**
TIMP1	**Tissue inhibitor of metalloproteinase 1**	**NM_003254**	**16[a]**
TRO	Trophinin/magphinin	AB029037	14[a]
UREB1	**Upstream regulatory element binding protein 1**	**AB002310**	**12[a,b]**
UXT	**Ubiquitously-expressed transcript**	**AF092737**	**18[b]**
WDR13	**WD repeat domain 13 protein**	**NM_017883**	**18[b]**
ZNF41	**Zinc-finger protein 41**	**NM_153380**	**19[a]**

[a]Not screened for all 22 patients owing to limited amount of DNA.
[b]Not screened in patients with confirmed disease-causing mutation in the PQBP1 gene.
[c]Exon 1 of PCSK1N not screened.
Bold: Successful Northern blot hybridization.

designed using either the 'Primer3'[15] or the 'Pride' software.[16] Exons longer than 200 bp were divided into overlapping fragments suitable for mutation detection by denaturing HPLC (see below). The primer sequences are available upon request.

PCR

In general, PCR amplifications were carried out in 50 μl reaction volumes containing 100 ng genomic DNA, 1 $\times$ supplied reaction buffer, 10 pmol of each primer, 200 μM dNTPs and 1 U *Taq* polymerase (Promega, Mannheim, Germany or Qiagen, Hilden, Germany). A touchdown PCR profile was used. Step1: 96°C for 3 min followed by 20 cycles (95°C for 30 s, 65°C for 30 s) with a decrement of 0.5°/cycle. Step2: 30 cycles (95°C for 30 s, 55°C for 30 s and 72°C for 30 s). The PCR was concluded by a 5 min extension at 72°C. Alternatively, a PCR profile consisting of an initial denaturation step at 96°C for 3 min followed by 30–40 cycles at 95°C for 30 s, primer sequence-dependent annealing temperature for 45 s and 72°C for 30 s, with a 5 min final extension period (72°C) has been used. The specificity and the amount of the amplified products were checked by agarose gel electrophoresis before further analysis.

Denaturing high-performance liquid chromatography analysis

PCR-amplified fragments were submitted to denaturing high-performance liquid chromatography analysis (DHPLC) (WAVE Nucleic Acid Fragment Analysis System Transgenomic Inc., San Jose, CA, USA). For DHPLC, PCR products were pooled pairwise and denatured at 95°C for 10 min, followed by gradual re-annealing to room temperature over 20 min to enhance heteroduplex formation. Eight microliters of pooled PCR product were then injected into the autosampler, separated through a DNASep HT Cartridge (Transgenomic Inc., San Jose, CA, USA), eluted using a linear acetonitrile gradient at a flow rate of 0.9 ml/min and detected by UV analysis at 260 nm. Optimal conditions for each injection (temperature, elution time and buffer composition) were determined using the WAVE Maker software (version 4.1.40, Transgenomic). Samples were analyzed at 2–4 different temperatures in order to detect sequence variants in different melting domains of the fragments. DHPLC conditions for individual fragments are available upon request.

Sequence analysis

Sequencing reactions were carried out for patient DNAs, which showed abnormal elution profiles in the DHPLC analysis. Before sequencing, the original PCR products were either purified using the Qiaquick PCR Purification Kit (Qiagen, Hilden, Germany) or they were directly sequenced in both directions using a 3100 Genetic Analyzer and Big Dye terminator chemistry (Applied Biosystems, Foster City, CA, USA). Sequence data were assembled and analyzed using the GAP4 Contig Editor.[17]

RNA isolation and Northern blot analysis

Total RNA was isolated from patient LCL by use of Trizol (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's recommendations. Fetal brain RNA was purchased from BD Biosciences (Palo Alto, CA, USA). Poly-A+ RNAs, obtained from 100 μg total RNA by using Dynabeads oligo-dT_{25} (Dynal Biotech, Oslo, Norway), were separated on a formaldehyde containing gel in 1 $\times$ MOPS, transferred to a Hybond N^+ membrane and crosslinked by UV using the auto-crosslink program of a Stratalinker (Stratagene, La Jolla, CA, USA). The gene-specific probes with an average size between 300 and 600 bp were PCR amplified from genomic DNA. All probes were designed to hybridize to at least 100 bases of the respective RefSeq cDNA. The specificity of the probes was checked by BLAST alignment. The sequences of primers used for probe generation are available upon request. Probes were labeled with ^{32}[P]dCTP using Klenow enzyme and random hexamer primers. The labeled fragments were purified and hybridized to membranes in UltraHyb buffer (Ambion, Austin, TX, USA) and washed according to instructions of the manufacturer. Subsequently, Northern blots were exposed to Fuji Medical X-Ray films at −80°C for 6 h up to 8 days or were analyzed using a Storm 820 imaging system (APBiotech, Piscataway, NJ, USA). To control for RNA loading, blots were re-probed with a β-actin probe (BioChain, Hayward, CA, USA).

Results

Spectrum of sequence variants

Within the framework of our molecular screen, we have studied a total of 47 genes in up to 22 different XLMR families. *PQBP1*, *FTSJ1* and *JARID1C* were screened in all 22 families, whereas 15 other genes were screened only in families where no *PQBP1* mutations had been found (indicated by '[a]' in Table 2). For the remaining genes, fewer families have been analyzed, because in six patients, only limited amounts of DNA were available (genes denoted with '[b]' in Table 2). With regard to the 22 MRX families analyzed, 17 did not carry mutations in *PQBP1* and 11 of these have been screened for all 46 remaining genes. Altogether, 705 different PCR fragments (covering 548 exons) that represent about 90 000 bp of coding and about 60 000 bp of noncoding DNA sequence have been screened for the majority of patients.

The distribution of the 57 variants is shown in Figure 1. The sequence variants are not evenly distributed among the 47 investigated genes: in 27 genes, one or several variants were found, whereas no variants were found for the remaining 20 genes.

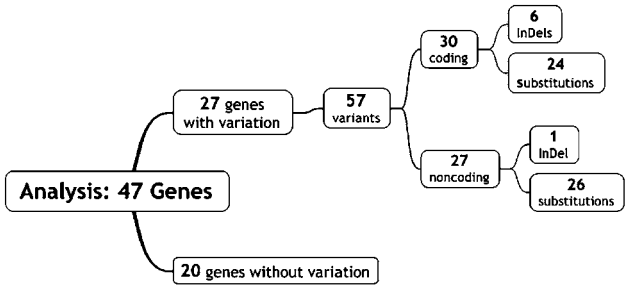

Figure 1 Distribution of sequence variants among the 47 investigated genes is shown. Only 27 genes were affected by one or more changes and the proportion of variants in coding and noncoding regions are nearly equal.

Table 3 Summary of sequence alterations identified in screening 47 candidate genes in Xp11 in a panel of 22 patients with mental retardation

Gene name	*Nucleotide change*	*Amino-acid variation*	*Number of patients with variation*
CCNB3	c.3424−42A>G	—	4
FGD1	c.659+27T>C	—	3
FLJ10613	c.631−18T>C	—	1
FLJ14103	**c.70T>C**	**S24P**	1
	c.383A>G	K128R	3
	c.486G>A	**P162P**	1
FLJ21687	c.−47C>T	—	3
	c.442C>T	**R148W**	1
FTSJ1	c.571+69C>T	—	2
	c.655G>A	**D219N**	1
GRIPAP1	c.607−48G>A	—	4
HDAC6	**c.513C>T**	**Y171Y**	1
	c.2495G>A	**R832H**	2
JARID1C	**c.522+19G>A**	—	2
	c.1162G>C	**A 388 P**	1
	c.2191C>T	**L 731 F**	1
JM1	c.51−25T>C	—	4
	c.361+17A>G	—	1
	c.747G>A	Q 249 Q	4
	c.1540−18T>C	—	2
KIAA1202	**c.2957+21T>C**	—	1
	c.4829A>G	—	6
	c.3350_3361del	Q1124_Q1127del	12
	c.3424_3426del	E1142del	12
	c.3483G>A	E1161E	12
MAGED2	c.252A>G	S84S	5
	c.624C>T	A208A	5
	c.981T>C	S327S	1
	c.1443A>G	A481A	1
MG61/PPN	c.373+150G>A	—	1
PCSK1N	**c.143G>A**	**E 39 E**	1
	c.450C>A	L 150 L	2
PHF8	c.1050delACAG gtcttccc	T351delinsX375	1
PLP2	**c.434A>C**	**T 145 N**	1
PQBP1	c.180−3C>T	—	4
	c.450_451insAG	R155fsXS184	2
	c.450-453delAG	E151fsXA163	1
	c.450-455delAGAG	R153fsXS182	1
RBM3	**c.−13−108C>T**	—	1
	c.−13−57C>G	—	1
	c.−13−27C>T	—	1
SLC35A2	c.1164−51G>A	—	1
SLC38A5	c.54−33T>C	—	5
	c.1352T>C	M 451 T	2
SMC1L1	c.−19C>T	—	1
	c.3585C>G	**T 1195 T**	1
T54	**c.1141−47C>T**	—	2
TFE3	**c.1004−43C>A**	—	3
	c.1323G>A	V441V	1
TRO	c.1525+68_69 delCT	—	1
	c.1316C>T	**P439L**	3
UREB1	c.579G>A	S193S	3
	c.2502+20G>A	—	1
	c.8056+25G>C	—	1
WDR13	c.832−210G>A	—	2
	c.1154+31G>C	—	1
ZNF41	**c.374T>G**	**I125R**	1

Bold: Not in dbSNP.

Of the 57 variants, 50 are nucleotide substitutions (40 transitions and 10 transversions, see Table 4), 21 of which have not been reported in dbSNP (marked in bold in Table 3).

Of the 30 sequence variants found in the coding region (Figure 1, Table 3), 10 are putatively pathogenic alterations affecting the following genes: *PQBP1*, *FTSJ1*, *PHF8*, *JARID1C*, *PLP2* and *FLJ14103*. In *FLJ14103*, a single variant (c.70G>T, S24P) was found, which was not present in 95 controls (with higher educational background) and 180 unrelated MR patients. The variant in *PLP2* (c.434A>C, T145N) was not found in >600 controls. The *FLJ14103* variant does not segregate with the disorder, whereas the *PLP2* variant cosegregates with MR in all families where it

Table 4 Summary of 57 sequence variants found in 47 genes

Sequence variants	*Coding*	*Noncoding*	*Total*
40 transitions			
A>G	3	3	6
C>T	5	7	12
G>A	8	6	14
T>C	2	6	8
10 transversions			
A>C	1	0	1
C>A	1	1	2
C>G	1	1	2
G>C	1	2	3
G>T	1	0	1
T>G	1	0	1
One insertion	1	0	1
Six deletions	5	1	6
Sum	31	27	57

has been observed. The eight mutations we found in the four remaining genes have already been reported elsewhere.[18–21] Interestingly, the variant c.442C>T (R148W) in *FLJ21687* was not detected in the control panel, but the analysis of 180 patients from small MR families revealed the same variation in two patients, one of them being from family P048 where a nonsense mutation in *FTSJ1* has been described.[18]

Five further missense variants (detected in the 22 MRX families) were found in at least two out of 250 control X-chromosomes. These variants include *FLJ14103* (K128R), *HDAC6* (R832H), *SLC38A5* (M451T), *TRO* (P439L) and *ZNF41* (I125R), of which the variants detected in *FLJ14103* and *SLC38A5* were both present in dbSNP.

Thirteen different silent sequence variants were encountered in nine different genes. Seven silent variants were found only once in our patient cohort, and four of these have not been described in dbSNP so far. All other silent variants were found more than once in our patient panel and were present in dbSNP. Other variants that have not been published or could not be found in controls or in dbSNP were not observed in the panel of 180 small MR families.

The most common silent variant (found in 12 out of 16 patients in our patient cohort) is the transition G>A at nucleotide 3483 (E1161E) in the gene *KIAA1202*. This transition is part of a haplotype, which also contains a 12 bp deletion c.3350_3361del (Q1124_Q1127del) and a deletion of 3 bp (c.3424_3426del; E1142del). Both haplotypes were found in controls and are described by Hagens *et al.*[22]

In MAGED2, another haplotype block was identified as five patients carry the transition c.252A>G in combination with c.624C>T. The other analyzed patients do not carry any of these variants.

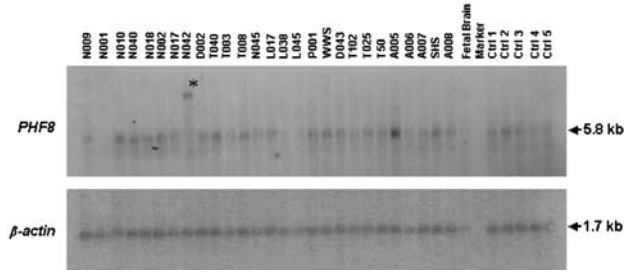

Figure 2 Northern blot analysis of poly-A+ RNA from 27 patient and five control LCL. RNA from fetal brain has been included as control. The blot was sequentially hybridized with a cDNA probe corresponding to the 3′UTR of *PHF8* and a *â-actin* cDNA probe. (*) Note the increased transcript size in the patient from family N042.

In noncoding regions, we have identified 27 different sequence variants (Figure 1) including one 2 bp deletion and 26 SNPs. The 2 bp deletion was found in the trophinin gene, 68 bp downstream of the donor splice site of exon 6, in the patient of family N017 (also carrying a missense mutation c.1162G>C, A388P in *JARID1C*). The majority of the 26 SNPs were present in dbSNP except for eight variants, four of which were found only once in the patients tested.

Expression analysis of candidate genes by Northern blot hybridization

Clinically relevant sequence variants are not necessarily confined to coding regions: changes in regulatory sequences can alter gene expression levels, and intronic mutations may result in altered splicing patterns. To detect such variants, we have carried out Northern blot analysis in 27 XLMR patients (including the 22 patients analyzed for sequence variation plus five more recently collected patients). Of the 47 candidate genes investigated, 26 yielded specific signals corresponding to the expected transcript size (denoted bold in Table 2). mRNA expression of the 26 genes was analyzed in five separate Northern blot experiments using membranes carrying the same amount of RNA from each patient. Results were assessed by visual comparison of signal intensity patterns in patients and controls. Low signal intensities were consistently observed for patients N009, N001, N017, T003, N045, L038 and L045. Hybridization signals from lymphoblastoid RNA and fetal brain RNA showed different intensities, but banding patterns were similar in both tissues. Northern blot hybridization of *PQBP1* confirmed our previous results: patients from families N009, N040, N045, MRX55 and SHS showed an almost complete loss of *PQBP1* expression,[20] and abnormal splicing of *FTSJ1* was observed for the patient of family MRX44, where a G>A transition had been found.[18] Abnormal splicing was also observed for the patient of family N042, where a 12 bp deletion covering

the donor splice site of *PHF8* exon 8 (c.1050delA-CAGgtcttccc) had been identified (Figure 2), which confirms the results recently published by Laumonnier *et al*[21] Some genes, including *APE2*, *HADH2*, *PLP2* and *TIMP1*, displayed relatively strong variation, both in patients and in controls, but we did not observe consistent mRNA expression changes pointing to effects of silent or intronic sequence variation.

Discussion

MR can result from selective impairment of brain development or physiology but also from fundamental cellular defects that are present in many tissues, but predominantly affecting the brain, for example, because of its higher sensitivity.[23] This explains why some MR genes are expressed specifically in the brain, whereas others are ubiquitously expressed. The prevalence of MR is lower in females than in males, which is partially due to mutations on the X-chromosome (for a detailed discussion on the subject, see Ropers and Hamel[2]). Based on the distribution of linkage intervals in families affected with MRX, we have previously shown that approximately 30% of the causative mutations localize to the proximal Xp and the pericentromeric region.[3] Within this area, a 7.4 Mb region flanked by the genes *ELK1* and *ALAS2* contains the highest number of defects that give rise to MRX.[3]

Therefore, we have selected 47 genes within this interval that are expressed in nervous tissue (but not necessarily exclusively so) and analyzed them for mutations in up to 22 MR families with linkage to this area. This led to the identification of four genes[18–21] involved in MRX, a comparatively high number, which underscores the heterogeneity of this disorder and at the same time confirms the *ELK1–ALAS2* region as a hotspot for MRX candidate genes.

In the remaining 43 genes analyzed in this screen, we found eight missense variants, three of which are not present in controls or in dbSNP, R148W in *FLJ21687*, T145N in *PLP2* and S24P in *FLJ14103*. The *FLJ21687* variant was not found in controls, but a small MR family, P048, carries this variant as well as a nonsense mutation in *FTSJ1*. Under the likely assumption that only one mutation is responsible for the disease in each family, R148W in *FLJ21687* is unlikely to be causative. For the patient in whom the variant in *PLP2* (T145N) was found, no other DNA changes have been reported. Therefore, this variant could be involved in MRX, but at present there is no functional evidence to support this.

Intronic variants can affect splicing and mRNA stability and increasing evidence suggests that this is also the case for silent variants (for a detailed discussion on the subject, see Chamary *et al*[24]). To study the possible presence of such variants, Northern blot analysis was carried out for 26 of the 47 investigated genes (the remaining 21 genes were not expressed at detectable levels in LCL). As expected, the frameshift mutations in *PQBP1* and the splice site mutations in *FTSJ1* and *PHF8* could be shown to alter mRNA expression or splicing in LCLs. As no other detectable effects on mRNA expression were observed for the known MRX genes *PQBP1*, *FTSJ1*, *PHF8*, *JARID1C* and *ZNF41*, this suggests that silent or non-coding variation in these genes is not a common cause of MRX.

In order to address the question whether brain tissues express specific splice variants of MR candidate genes that might not be detectable in LCLs, we included fetal brain RNA in our Northern blot analysis. As we could not observe differential splice patterns between LCLs and fetal brain for the 26 genes where Northern blotting was successful, we conclude that none of the major transcripts escaped our LCL-based analysis. Still, expression of 21 genes in LCLs was too low for detection using Northern blot hybridization, and in seven of these genes (*FLJ14103*, *FLJ21687*, *KIAA1202*, *PCSK1N*, *RBM3*, *SMC1L1* and *TRO*) we found nine sequence variants that were not present in dbSNP. The possibility that (some of) these variants have an effect on mRNA expression or splicing cannot be excluded.

Our findings also comprise a number of small (<13 bp) deletions. Owing to the limitations of our approach (PCR and sequencing), larger genomic rearrangements (except deletions) could not be detected in this study. Duplications containing *MECP2*[25] have been described as frequent cause of Rett syndrome, a syndromic form of MRX, and it is conceivable that similar rearrangements occur elsewhere on the X, too. However, low copy repeats, which often mediate duplications or inversions are comparatively rare in the Xp11 region,[26] suggesting that in this region, large genomic rearrangements are not common. This also fits with the general absence of aberrant mRNA expression patterns, except for changes that could be ascribed to abnormal mRNA processing due to known mutations.

Taken together, in this study, we have found mutations within four genes, in eight out of 22 families with linkage to the *ELK1–ALAS2* region. The majority of mutations in the 22 families are not affecting the five MRX genes *PQBP1*, *FTSJ1*, *PHF8*, *JARID1C* and *ZNF41*.[18–21,27] Therefore, it is very likely that mutation analysis in other families with linkage intervals overlapping this region will show a similarly low proportion of mutations affecting the above-mentioned genes, which implies that the majority of mutations is harbored by other genes.

As our study has excluded MRX causative mutations in 43 of these other genes in nine families with linkage information, their potential of bearing relevant changes for X-linked MR has been demonstrated to be low. This information will prove valuable when prioritizing candidate genes in the search for mutations in further MR families.

Acknowledgements

We are grateful to Drs Robledo and Siniscalco for providing patient material of family MRX26 and to Marion Amende, Kirsten Hoffmann as well as Melanie Wendehack for excellent technical support. We also thank Ulf Gurok, Sebastian Haesler, Beatriz Aranda, Nils Hartmann, Vera Kalscheuer, Sarah Shoichet, Olivier Hagens, Jiong Tao, Stefan Haas, and Alfred Beck. The study was supported by the EURO-MRX project, part of the European 'Quality of Life and Management of Living Resources Programme' QLG3-CT-2002-01810.

Electronic-database information/URLs

BLAST2: http://www.ncbi.nlm.nih.gov/BLAST/
Genome Browser: http://genome.cse.ucsc.edu/cgi-bin/hgGateway?org=human
NCBI: http://www.ncbi.nlm.nih.gov/
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Pride: http://pride.molgen.mpg.de/genomepride.html
Primer3: http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
Unigene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene

References

1 Roeleveld N, Zielhuis GA, Gabreels F: The prevalence of mental retardation: a critical review of recent literature. *Dev Med Child Neurol* 1997; **39**: 125–132.
2 Ropers HH, Hamel BC: X-linked mental retardation. *Nat Rev Genet* 2005; **6**: 46–57.
3 Ropers HH, Hoeltzenbein M, Kalscheuer V *et al*: Nonsyndromic X-linked mental retardation: where are the missing mutations? *Trends Genet* 2003; **19**: 316–320.
4 Sutherland GR, Gedeon AK, Haan EA, Woodroffe P, Mulley JC: Linkage studies with the gene for an X-linked syndrome of mental retardation, microcephaly and spastic diplegia (MRX2). *Am J Med Genet* 1988; **30**: 493–508.
5 Kloos DU, Jakubiczka S, Wienker T, Wolff G, Wieacker P: Localization of the gene for Wieacker–Wolff syndrome in the pericentromeric region of the X chromosome. *Hum Genet* 1997; **100**: 426–430.
6 Raynaud M, Gendrot C, Dessay B *et al*: X-linked mental retardation with neonatal hypotonia in a French family (MRX15): gene assignment to Xp11.22–Xp21.1. *Am J Med Genet* 1996; **64**: 97–106.
7 Gedeon A, Kerr B, Mulley J, Turner G: Pericentromeric genes for non-specific X-linked mental retardation (MRX). *Am J Med Genet* 1994; **51**: 553–564.
8 Robledo R, Melis P, Laficara F *et al*: Further linkage evidence for localization of mutational sites for nonsyndromic types of X-linked mental retardation at the pericentromeric region. *Am J Med Genet* 1996; **64**: 107–112.
9 Hamel BC, Smits AP, van den Helm B *et al*: Four families (MRX43, MRX44, MRX45, MRX52) with nonspecific X-linked mental retardation: clinical and psychometric data and results of linkage analysis. *Am J Med Genet* 1999; **85**: 290–304.
10 Deqaqi SC, N'Guessan M, Forner J *et al*: A gene for non-specific X-linked mental retardation (MRX55) is located in Xp11. *Ann Genet* 1998; **41**: 11–16.
11 Yntema HG, van den Helm B, Knoers NV *et al*: X-linked mental retardation: evidence for a recent mutation in a five-generation family (MRX65) linked to the pericentromeric region. *Am J Med Genet* 1999; **85**: 305–308.
12 Raynaud M, Moizard MP, Dessay B *et al*: Systematic analysis of X-inactivation in 19XLMR families: extremely skewed profiles in carriers in three families. *Eur J Hum Genet* 2000; **8**: 253–258.
13 Claes S, Devriendt K, Van Goethem G *et al*: Novel syndromic form of X-linked complicated spastic paraplegia. *Am J Med Genet* 2000; **94**: 1–4.
14 Siderius LE, Hamel BC, van Bokhoven H *et al*: X-linked mental retardation associated with cleft lip/palate maps to Xp11.3–q21.3. *Am J Med Genet* 1999; **85**: 216–220.
15 Rozen S, Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000; **132**: 365–386.
16 Haas S, Vingron M, Poustka A, Wiemann S: Primer design for large scale sequencing. *Nucleic Acids Res* 1998; **26**: 3006–3012.
17 Bonfield JK, Smith K, Staden R: A new DNA sequence assembly program. *Nucleic Acids Res* 1995; **23**: 4992–4999.
18 Freude K, Hoffmann K, Jensen LR *et al*: Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. *Am J Hum Genet* 2004; **75**: 305–309.
19 Jensen LR, Amende M, Gurok U *et al*: Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *Am J Hum Genet* 2005; **76**: 227–236.
20 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.
21 Laumonnier F, Holbert S, Ronce N *et al*: Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *J Med Genet* 2005; **42**: 780–786.
22 Hagens O, Dubos A, Abidi F *et al*: Disruptions of the novel KIAA1202 gene are associated with X-linked mental retardation. *Hum Genet* 2006; **118**: 578–590.
23 Inlow JK, Restifo LL: Molecular and comparative genetics of mental retardation. *Genetics* 2004; **166**: 835–881.
24 Chamary JV, Parmley JL, Hurst LD: Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat Rev Genet* 2006; **7**: 98–108.
25 Van Esch H, Bauters M, Ignatius J *et al*: Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 2005; **77**: 442–453.
26 She X, Horvath JE, Jiang Z *et al*: The structure and evolution of centromeric transition regions within the human genome. *Nature* 2004; **430**: 857–864.
27 Shoichet SA, Hoffmann K, Menzel C *et al*: Mutations in the ZNF41 gene are associated with cognitive deficits: identification of a new candidate for X-linked mental retardation. *Am J Hum Genet* 2003; **73**: 1341–1354.