

ARTICLE

Next-generation sequencing in X-linked intellectual disability

Andreas Tzschach^{*,1,2,14}, Ute Grasshoff^{2,14}, Stefanie Beck-Woedl^{2,14}, Claudia Dufke², Claudia Bauer², Martin Kehrer², Christina Evers³, Ute Moog³, Barbara Oehl-Jaschkowitz⁴, Nataliya Di Donato¹, Robert Maiwald⁵, Christine Jung⁶, Alma Kuechler⁷, Solveig Schulz⁸, Peter Meinecke⁹, Stephanie Spranger¹⁰, Jürgen Kohlhasse¹¹, Jörg Seidel¹², Silke Reif¹³, Manuela Rieger¹³, Angelika Riess², Marc Sturm², Julia Bickmann², Christopher Schroeder², Andreas Dufke², Olaf Riess² and Peter Bauer²

X-linked intellectual disability (XLID) is a genetically heterogeneous disorder with more than 100 genes known to date. Most genes are responsible for a small proportion of patients only, which has hitherto hampered the systematic screening of large patient cohorts. We performed targeted enrichment and next-generation sequencing of 107 XLID genes in a cohort of 150 male patients. Hundred patients had sporadic intellectual disability, and 50 patients had a family history suggestive of XLID. We also analysed a sporadic female patient with severe ID and epilepsy because she had strongly skewed X-inactivation. Target enrichment and high parallel sequencing allowed a diagnostic coverage of > 10 reads for ~96% of all coding bases of the XLID genes at a mean coverage of 124 reads. We found 18 pathogenic variants in 13 XLID genes (*AP1S2*, *ATRX*, *CUL4B*, *DLG3*, *IQSEC2*, *KDM5C*, *MED12*, *OPHN1*, *SLC9A6*, *SMC1A*, *UBE2A*, *UPF3B* and *ZDHHC9*) among the 150 male patients. Thirteen pathogenic variants were present in the group of 50 familial patients (26%), and 5 pathogenic variants among the 100 sporadic patients (5%). Systematic gene dosage analysis for low coverage exons detected one pathogenic hemizygous deletion. An *IQSEC2* nonsense variant was detected in the female ID patient, providing further evidence for a role of this gene in encephalopathy in females. Skewed X-inactivation was more frequently observed in mothers with pathogenic variants compared with those without known X-linked defects. The mutation rate in the cohort of sporadic patients corroborates previous estimates of 5–10% for X-chromosomal defects in male ID patients.

European Journal of Human Genetics (2015) 23, 1513–1518; doi:10.1038/ejhg.2015.5; published online 4 February 2015

INTRODUCTION

Intellectual disability (ID) is a childhood-onset neurodevelopmental disorder commonly defined by an intelligence quotient below 70.¹ It has an estimated incidence of 1–3% in Western societies.² Causes of ID include environmental factors and genetic defects. Mutations in X-chromosomal genes are estimated to account for 5–10% of male ID patients.³ In female carriers of defects in X-chromosomal genes, inactivation of the X-chromosomes can be skewed (ie, non-random), which may be used as a predictive marker for such mutations.⁴ X-linked intellectual disability (XLID) is characterized by extensive genetic heterogeneity; to date, mutations in more than 100 genes on the X-chromosome are known to be associated with XLID.⁵ With the exception of Fragile X syndrome that is at cause in about 25% of XLID patients, the other XLID genes individually only account for a small percentage of patients.⁶ This has rendered routine testing of these genes impractical in patients with nonspecific clinical features. Next-generation sequencing (NGS) technologies have recently provided the means for overcoming this problem by cost-effective parallel analysis of large numbers of genes. In contrast to the use of NGS as a research

tool for which reduced sensitivity can be acceptable, application of NGS in diagnostics demands for more rigid quality standards, as insufficient coverage and/or inaccurate bioinformatic analysis will lower the diagnostic yield. In order to maximize sensitivity and specificity, targeted gene panels make use of high coverage sequencing at relatively low costs.⁷

Here, we report on the screening of 107 XLID genes that were analysed by targeted enrichment and subsequent NGS in a cohort of 150 male ID patients.

PATIENTS AND METHODS

Patients

The study comprised 150 male ID patients in whom the cause of the disease was not known. One hundred patients were sporadic cases, that is, they had no other family members with cognitive defects. Fifty patients had a family history suggestive of XLID, that is, they had either affected brothers or male maternal relatives or less severely affected female family members compatible with an X-linked mode of inheritance.

Chromosome aberrations were ruled out by microarray analysis (array CGH or SNP array) in all patients before they were included into this study. Fragile X

¹Institute of Clinical Genetics, Technische Universität Dresden, Dresden, Germany; ²Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany; ³Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany; ⁴Practice of Human Genetics, Homburg (Saar), Germany; ⁵Medical Care Center, Mönchengladbach, Germany; ⁶Centre of Human Genetics, Mannheim, Germany; ⁷Institute of Human Genetics, University Hospital Essen, Essen, Germany; ⁸Center of Human Genetics, Jena University Hospital, Jena, Germany; ⁹Institute of Human Genetics, University Hospital Hamburg-Eppendorf, Hamburg, Germany; ¹⁰Practice of Human Genetics, Bremen, Germany; ¹¹Centre for Human Genetics, Freiburg, Germany; ¹²Practice of Pediatrics, Jena, Germany; ¹³Practice of Medical Genetics, Dresden, Germany
¹⁴These authors contributed equally to this work.

*Correspondence: Professor A Tzschach, Institute of Clinical Genetics, Technische Universität Dresden, Fetscherstr. 74, Dresden 01307, Germany. E-mail: andreas.tzschach@uniklinikum-dresden.de; Tel: +49 351 458 6817; Fax: +49 351 458 4316

Received 18 March 2014; revised 2 January 2014; accepted 7 January 2015; published online 4 February 2015

syndrome was excluded in all patients by testing for CGG repeat expansions in the 5' untranslated region of *FMR1*. Patients with a history of seizures were also tested for the recurrent 24 bp duplication in exon 2 of *ARX*.

All patients were seen by a clinical geneticist to rule out obvious syndromic disorders. Four patients were excluded in this way from the main cohort because recognizable syndromes (Allan-Herndon-Dudley syndrome, Börjeson-Forsman-Lehmann syndrome and Christianson syndrome) were suspected and subsequently confirmed by conventional Sanger sequencing of the respective genes. We also excluded patients with indications for autosomal dominant disorders (eg, if they had a similarly affected father) or autosomal recessive defects (eg, if they had similarly affected sisters and consanguineous parents).

In addition to the 150 male ID patients of the main cohort, we also analysed a female sporadic patient with severe ID and epilepsy because she had strongly skewed X-inactivation (patient 19, Table 1 and Supplement).

Next-generation sequencing

DNA samples from peripheral blood cells were prepared according to standard protocols. Fragmentation of patient DNA samples was carried out with a Covaris S2 ultrasonicator, generating fragments of 200–500 bp in size (Covaris, Woburn, MA, USA). For library generation of the fragmented samples, the Illumina Paired-End DNA Sample Preparation Kit was used (Illumina, San Diego, CA, USA).

Enrichment of the target sequences, which included nearly all coding and flanking intronic regions of the X-chromosome, was performed using the Agilent SureSelect^{XT} X-Chromosome in-solution target enrichment kit (Agilent, Santa Clara, CA, USA) followed by individual barcoding of each sample.

Sequencing was performed using the Illumina GAIIx sequencer (2 × 76 paired-end sequencing). For each lane of a flow cell, six patient samples were pooled.

The 76 bp reads were preprocessed with adapter trimming (cutadapt)⁸ duplicate removal (picard-tools version 1.75; <http://picard.sourceforge.net>). The reads were subsequently mapped to the human reference genome (hg19) using BWA (version 0.5.9) and Stampy (version 1.0.17).^{9,10} Variants were detected using SAMtools (version 0.1.18) and annotated with data from the 1000 Genomes Project, dbSNP and the NHLBI GO Exome Sequencing Project using ANNOVAR (version 2012-05-25).^{11,12}

For gene dosage analysis (hemizygous deletions in males), a simple but effective statistical approach was used. First, the normalized average read depth of each exon was calculated (ie, the average read depth of an exon divided by the average read depth of all exons of the sample). Then, the mean and standard deviation (SD) of the normalized average read depth over all samples were calculated for each exon. All normalized read depth values outside a three SD window around the mean were PCR-validated after apparent false positives had been excluded by inspection in the Integrative Genomics Viewer (IGV; <http://www.broadinstitute.org/igv/>).

Analysis of variants was restricted to the coding and flanking intronic sequences of the following 107 XLID genes: *ACSL4*, *AFF2*, *AGTR2*, *AIFM1*, *AP1S2*, *ARHGEF6*, *ARHGEF9*, *ARX*, *ATP6AP2*, *ATP7A*, *ATRX*, *BCOR*, *BRWD3*, *CASK*, *CDKL5*, *CLCN4*, *CUL4B*, *DCX*, *DKC1*, *DLG3*, *DMD*, *EIF2S3*, *FANCB*, *FGD1*, *FLNA*, *FMR1*, *FTSJ1*, *GDI1*, *GK*, *GPC3*, *GRIA3*, *HCCS*, *HCFC1*, *HDAC8*, *HPRT*, *HSD17B10*, *HUWE1*, *IDS*, *IGBP1*, *IKBK*, *IL1RAPL1*, *IQSEC2*, *KDM5C*, *KDM6A*, *KIAA2022*, *KLF8*, *LICAM*, *LAMP2*, *LASIL*, *MAGT1*, *MAOA*, *MBTPS2*, *MECP2*, *MED12*, *MID1*, *MTM1*, *NAA10*, *NDP*, *NDUFA1*, *NHS*, *NLGN3*, *NLGN4X*, *NSDHL*, *NXF5*, *OCRL*, *OFD1*, *OPHN1*, *OTC*, *PAK3*, *PCDH19*, *PDHA1*, *PGK1*, *PHF6*, *PHF8*, *PLP1*, *PORCN*, *PQBPI*, *PRPS1*, *PTCHD1*, *RAB39B*, *RAB40AL*, *RBM10*, *RPL10*, *RPS6KA3*, *SHROOM4*, *SLC6A8*, *SLC9A6*, *SLC16A2*, *SMC1A*, *SMS*, *SOX3*, *SRPX2*, *SYP*, *SYN1*, *THOC2*, *TIMM8A*, *TSPAN7*, *UBE2A*, *UPF3B*, *WDR45*, *ZDHHC9*, *ZDHHC15*, *ZMYM3*, *ZNF41*, *ZNF81*, *ZNF674* and *ZNF711*. References for all genes are in Supplementary Table S1 of the review by Lubs *et al*,³ except for *SYP*, *ZNF711*, *AIFM1*, *ZMYM3*, *KDM6A* and *WDR45*.^{13–18}

Variant classification

All coding variants and variants in the exon-intron boundaries (+/–10 bp) were classified following the suggestions of Plon *et al*.¹⁹ Only probably (VUS4)

or definitely pathogenic variants (VUS5) were used to establish a genetic diagnosis. Novel missense variants, synonymous variants with potential effects on splicing and in-frame deletions or insertions were rated VUS3 (uncertain significance) regardless of *in silico* predictions. A novel missense variant was assigned as probably pathogenic (VUS4) if segregation was informative and the clinical phenotype met published expectations. Nonsense, frameshift and splicing variants (+/–2 bp) were also classified as VUS4 if they were novel.

Sanger sequencing

Putatively pathogenic variants were validated by conventional Sanger sequencing. This was followed by segregation analysis of confirmed variants in affected and unaffected members of the respective families.

X-inactivation analysis

A total of 107 maternal DNA samples were available for X-inactivation analyses. We used the polymorphic CAG_n repeat within the human *AR* gene to assess the relative methylation status of both chromosomes after methylation-sensitive restriction enzyme digest.²⁰ X-inactivation was considered as skewed (non-random) if the ratio of the two alleles exceeded 80:20.

RESULTS

NGS sequencing characteristics

On average, 9.26 million reads of 76 bp length were generated per patient sample. A minimum 10-fold coverage per base was on average achieved for 94.61% of the target region. Regions with insufficient coverage were mostly recurrent and usually characterized by a high GC content. Specifically, parts of the coding regions of *IKBGK*, *IQSEC2*, *ARX*, *MBTBS2*, *PCDH19*, *SLC16A2*, *SOX3* and *ZNF81* could not be analysed reliably. *ABCD1* was altogether excluded from this study because sequence homology to other parts of the genome hampered analysis by this approach. Repetitive DNA structures such as trinucleotide DNA repeats could not be analysed with this method. In particular, this concerned the CGG-trinucleotide stretch in the 5' UTR of *FMR1*, expansion of which had been excluded in all patients before they were subjected to analysis by this panel. In patients with epilepsy, the recurrent 24 bp duplication in exon 2 of *ARX* was excluded by separate tests. The data were also analysed for sequencing drop-outs of entire exons which would suggest larger deletions. In one patient (patient 16, Table 1 and Supplement), a deletion of a single exon (exon 1 of *SLC9A6*) was detected in this way and confirmed by PCR. In contrast, the presence of larger duplications could not be excluded. Duplications larger than 100 kb would, however, have been detected by the array-based chromosome tests that had been performed in all patients before they were included in this study.

Variants with a minor allele frequency of more than 1%, intronic variants with a distance of more than 10 nucleotides to the nearest exon and synonymous variants without a predicted splicing effect were excluded from further analysis.

Pathogenic variants

In total, 18 probably or definitely pathogenic variants in 13 XLID genes were detected in the 150 patients of the main cohort (Table 1 and Supplement).

In the group of 50 patients with a family history suggestive of XLID, 13 pathogenic variants (26%) were detected in the following genes: *AP1S2* (family 1), *ATRX* (families 2, 3 and 4), *CUL4B* (families 5 and 6), *IQSEC2* (family 7), *KDM5C* (families 8 and 9), *MED12* (family 10), *OPHN1* (family 11), *UPF3B* (family 12) and *ZDHHC9* (family 13).

Among the 100 sporadic patients, five pathogenic variants (5%) were detected in *CUL4B* (patient 14), *DLG3* (patient 15), *SLC9A6* (patient 16), *SMC1A* (patient 17) and *UBE2A* (patient 18).

Table 1 Mutations identified in XLID patients

Patient	Gene	Variant	Interpretation	LOVD variant ID
<i>Patients with a family history suggestive of XLID</i>				
1	AP1S2	NM_003916.3: c.367C>T; p.(Gln123*)	nonsense variant; also present in mother and maternal grandmother	36859
2	ATRX	NM_000489.3: c.6975+5G>A; r.6850_6975del	variant segregates with ID in the family; skipping of exon 32 (125 bp) confirmed by RT-PCR studies	36706
3	ATRX	NM_000489.3: c.6244G>A; p.(Asp2082Asn)	PolyPhen2: probably damaging (score:1.000) SIFT: deleterious (score:0) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 23 highly conserved nucleotide (phyloP: 5.21) highly conserved amino acid segregation with ID in the family	36708
4	ATRX	NM_000489.3: c.7264_7265insT; p.(Gln2422Leufs*29)	frameshift mutation; segregation with ID	36846
5	CUL4B	NM_003588.3: c.2469C>G; p.(Ile823Met)	PolyPhen2: probably damaging (score:1.000) SIFT: deleterious (score:0.01) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 10 weakly conserved nucleotide (phyloP: 1.34) highly conserved amino acid segregation with ID in the family	36889
6	CUL4B	NM_003588.3: c.2060G>A; p.(Trp687*)	nonsense variant; segregation with ID in the family	36890
7	IQSEC2	NM_001111125.2: c.2662dup; p.(Ile888Asnfs*16)	frameshift variant (<i>de novo</i>)	36898
8	KDM5C	NM_004187.3: c.2047del; p.(Ala683Profs*81)	frameshift variant; segregation with ID in the family	36904
9	KDM5C	NM_001146702.1: c.3084C>A; p.(Cys1028*)	nonsense variant; segregation with ID in the family	36905
10	MED12	NM_005120.2: c.2444G>A; p.(Arg815Gln)	PolyPhen2: probably damaging (score:1.000) SIFT: deleterious (score:0) MutationTaster: disease causing (<i>P</i> -value: 1) Grantham distance: 43 highly conserved nucleotide (phyloP: 5.45) highly conserved amino acid segregation with ID in the family	36920
11	OPHN1	NM_002547.2: c.1489C>T; p.(Arg497*)	nonsense variant; segregation with ID in the family	36932
12	UPF3B	NM_080632.2: c.1101G>C; p.(Lys367Asn)	PolyPhen2: probably damaging (score:1.000) SIFT: tolerated (score:0.3) MutationTaster: disease causing (<i>P</i> -value:0.861) Grantham distance: 94 weakly conserved nucleotide (phyloP: 0.61) moderately conserved amino acid segregation with ID in the family	36963
13	ZDHHC9	NM_016032.3: c.286C>T; p.(Arg96Trp)	PolyPhen2: probably damaging (score 1.000) SIFT: deleterious (score:0) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 101 weakly conserved nucleotide (phyloP: 1.98) Highly conserved amino acid segregation with ID in the family	36967

Table 1 (Continued)

Patient	Gene	Variant	Interpretation	LOVD variant ID
<i>Sporadic patients</i>				
14	<i>CUL4B</i>	NM_003588.3: c.429_431dup; p.(Ser146dup)	in-frame insertion of one amino acid (serine); clinical features compatible with <i>CUL4B</i> mutation phenotype	36891
15	<i>DLG3</i>	NM_021120.3: c.649C>T; p.(Arg217*)	nonsense variant	36894
16	<i>SLC9A6</i>	NG_017160.1: g.135067656_135067991del (GRCh37/hg19)	deletion of 336 bp including exon 1 of <i>SLC9A6</i>	36955
17	<i>SMC1A</i>	NM_006306.2: c.1937T>C; p.(Phe646Ser)	PolyPhen2: probably damaging (score:1.000) SIFT: deleterious (score:0) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 155 moderately conserved nucleotide (phyloP: 2.87) highly conserved amino acid clinical features compatible with <i>SMC1A</i> mutation phenotype	SMC1A_00033
18	<i>UBE2A</i>	NM_003336.2: c.387dup; p.(Tyr130Valfs*9)	frameshift variant	36958
<i>Female ID patient with skewed X-inactivation</i>				
19	<i>IQSEC2</i>	NM_001111125.2: c.3163C>T; p.(Arg1055*)	frameshift variant	36899
<i>Patients in whom a specific syndrome was clinically suspected</i>				
20	<i>SLC16A2</i>	NM_006517.4: c.590G>A; p.(Arg197His)	PolyPhen2: probably damaging (score:1.000) SIFT: tolerated (score:0.07) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 29 moderately conserved nucleotide (phyloP: 2.47) moderately conserved amino acid identical variant already reported in an unrelated patient [HGMD: CM067032; published as NM_006517.3: c.812G>A; p.(Arg271His)]	36936
21	<i>PHF6</i>	NM_032458.2: c.687T>A; p.(His229Gln)	PolyPhen2: probably damaging (score:1.000) SIFT: deleterious (score:0) MutationTaster: disease causing (<i>P</i> -value:1) Grantham distance: 24 moderately conserved nucleotide (phyloP: 2.79) highly conserved amino acid variant also present in mother and sister	36728

LOVD, Leiden Open Variant database, www.lovd.nl.

In the female ID patient with skewed X-inactivation (patient 19), a nonsense variant was detected in *IQSEC2* [NM_001111125.2: c.3163C>T; p.(Arg1055*)] (Table 1 and Supplement).

Sanger sequencing in the male patients with a clinically diagnosed syndrome revealed a pathogenic variant in *SLC16A2* [NM_006517.4: c.590G>A; p.(Arg197His)]; according to the previous version of the reference sequence (NM_006517.3), this was c.812G>A; p.(Arg271His)] in patient 20 and a pathogenic variant in *PHF6* [NM_032458.2: c.687T>A; p.(His229Gln)] in patient 21 (Table 1 and Supplement).

The variants listed in Table 1 were submitted to the Leiden Open Variation Databases (www.lovd.nl).

Variants of uncertain significance

We also identified 42 rare variants (missense, synonymous or in-frame deletions) which did not qualify as benign polymorphisms (Supplementary Table 1). A pathogenic role for these variants was uncertain because there were either no family members available for segregation analyses, or the clinical features did not fit those that would have been expected for the respective gene, or the *in silico* prediction scores argued against a severe disruption of the gene product. With the availability of larger databases and/or functional studies, we expect that many (or most) of these variants will turn out to be benign, but some variants might eventually reveal themselves as causative defects.

X-inactivation in mothers

X-inactivation analysis was informative in mothers of 95 families, that is, two alleles of different size were present. Skewing was detected in 23 out of 95 mothers (24.2%); details of the ratios of skewed versus random X-inactivation in subgroups (families with pathogenic variants or VUS3 variants, families with sporadic patients or multiple affected members) are listed in Table 2. In families with pathogenic variants, skewing was more frequent (five of eight families, 62.5%) compared with the general cohort (24.2%), but a considerable part of this group (three out of eight families, 37.5%) had random X-inactivation (Table 2). Some female family members with skewed X-inactivation were also mildly retarded (eg, in family 9 with a pathogenic variant in *KDM5C*).

DISCUSSION

In this study, we have analysed a panel of 107 XLID genes by NGS. Outstanding sequence coverage was achieved with approximately 95% of all coding bases covered at >10 reads. As males are hemizygotous for X-linked genotypes, this coverage should have been sufficient to detect >99% of all covered SNVs.

In a cohort of 150 male ID patients, we have identified 18 pathogenic variants in these 107 genes. As expected, the diagnostic yield was higher among the 50 patients with a family history suggestive of X-chromosomal inheritance (26%) compared with the 100 sporadic patients (5%).

Bearing in mind that our cohort did not represent an unbiased sample of ID patients, these results nevertheless allow us to draw conclusions on the collective contribution of X-chromosomal genes to ID, and on the prevalence of mutations in individual XLID genes.

Fragile X syndrome, the most frequent single cause of XLID that is estimated to account for 25%, and submicroscopic copy number variations, which have been reported to be at cause in up to 10% of XLID patients, had been excluded in all patients before they were subjected to NGS.^{6,21} In patients with recognizable syndromes, the respective genes were directly tested by Sanger sequencing. This was the case in patient 20 with Allan-Herndon-Dudley syndrome, in patient 21 with Börjeson-Forsman-Lehmann syndrome (Table 1 and Supplement), and in two families with Christianson syndrome.²² Recognition of a syndrome does, however, rely on the clarity of the phenotypic manifestation and on the experience of the clinical geneticist. To avoid duplicate analyses, we refrained from direct gene testing unless the patients had unambiguous indications for a specific syndrome.

If one takes into consideration that (i) a proportion of male ID patients had been filtered out before by the criteria discussed above, (ii) mutations outside the immediate vicinity of the coding regions could not be detected, (iii) only approximately 95% of the target regions could be reliably analysed, (iv) some of the variants of

uncertain significance may in fact be pathogenic, and (v) that there may exist as yet unidentified additional XLID genes, the detection of five mutations in 100 sporadic patients is in line with previous estimates that mutations in X-chromosomal genes account for 5–10% of ID in males.³

With the exception of *ATRX* (pathogenic variants in three families), *CUL4B* (three pathogenic variants) and *KDM5C* (two pathogenic variants), only a single pathogenic variant was identified in each of the other 10 genes (*AP1S2*, *IQSEC2*, *MED12*, *OPHN1*, *UPF3B*, *ZDHHC9*, *DLG3*, *SLC9A6*, *SMC1A*, *UBE2A*) (Table 1). Six of those genes (*AP1S2*, *IQSEC2*, *MED12*, *UPF3B*, *ZDHHC9*, *DLG3*) have rarely been associated with ID, that is, fewer than 10 pathogenic variants have so far been reported for each individual gene.

The XLID-causing role of several of the 107 genes, which we analysed in this study, had been challenged in a recent re-evaluation based on data from large-scale exome sequencing projects.⁵ In none of the 17 genes that were put into question in this study and which were also included in our analysis (*ATP6AP2*, *ZNF674*, *ZNF41*, *ZNF81*, *MAGT1*, *SRPX2*, *NXF5*, *AGTR2*, *MAOA*, *SHROOM4*, *KLF8*, *IGBP1*, *NLGN3*, *ZMYM3*, *KIAA2022*, *ZDHC15*, *NAA10*), we detected pathogenic variants in our cohort of 150 ID patients.

Apart from broadening the mutational spectrum of rare XLID genes, the patients reported here also expand the clinical features associated with the respective genes. For example, gall bladder agenesis (as in the two brothers of family 4) is possibly an under-appreciated feature in patients with *ATRX* mutations. In patients with *CUL4B* mutations, neither cleft palate nor optic atrophy (as in family 5) nor microcephaly (as in patient 14) has been reported so far. Microcephaly is also an unusual feature in patients with *MED12* mutations (family 10).

XLID is not necessarily restricted to male mutation carriers. Among the families of the present study, this is illustrated by the affected women in family 9 (*KDM5C* mutation) who remind us of taking disease expression in females into consideration for genetic counselling.

The interpretation of intronic variants is a particular challenge. We have encountered this problem in family 2 with an intronic base pair exchange (c.6975+5G>A) in *ATRX*. Its pathogenic nature could only be established after segregation analyses and RT-PCR studies.

Egg donation is legally restricted in many countries, albeit not for genetic reasons. Patient 16, who suffered from Christianson syndrome, is an example for the genetic risks that can be associated with egg donation if the donor carries a pathogenic variant in an X-chromosomal gene.

We had also tested a sporadic female patient in whom strongly skewed X-inactivation hinted at an X-chromosomal disorder (patient 19, Table 1 and Supplement). This 16-year-old patient had severe ID, epilepsy and borderline macrocephaly. Her parents

Table 2 Results of X-inactivation studies in informative mothers

Family status	Random X-inactivation	Skewed X-inactivation	Total
All mothers	72 (75.8%)	23 (24.2%)	95
Mothers with positive family history	16 (73%)	6 (27%)	22
Mothers of sporadic patients (ie, with negative family history)	57 (78%)	16 (22%)	73
Mothers with VUS3 variants	21 (78%)	6 (22%)	27
All mothers with pathogenic variant	3 (37.5%)	5 (62.5%)	8
Mothers with positive family history and pathogenic variant	2 (33%)	4 (67%)	6
Mothers of sporadic patients (ie, with negative family history) and pathogenic variant	1 (50%)	1 (50%)	2

were healthy, and she had no other family members with ID or epilepsy. Previous genetic tests revealed normal results (array-CGH, FraX, *MECP2*, *CDKL5*). A stop mutation was detected in *IQSEC2* (Table 1). *IQSEC2* mutations had originally been identified in male ID patients (as in family 7 in the present report), and carrier females in the respective families were healthy.^{23,24} Only recently, a *de novo* stop mutation in *IQSEC2* was reported in a 3-year-old girl with epileptic encephalopathy.²⁵ The severity of the clinical problems of the female patient reported here is similar to that of the male patients with truncating *IQSEC2* mutations and may be explicable by skewing in favour of the X-chromosome with the mutated allele. However, even if this assumption is true, the cause of such disadvantageous skewing is presently not clear. A defect in a crucial gene on the other X-chromosome might provide a possible explanation. The female patient presented here, together with the girl that was reported by the Epi4K Consortium²⁵ extend the clinical manifestations of *IQSEC2* mutations beyond the male sex and should be considered in females with otherwise unexplained ID and epilepsy.

X-inactivation analyses in mothers of patients of our cohort have confirmed previous observations that skewing can be an indicator for an X-linked defect, but random X-inactivation (which was found in three out of eight informative carrier mothers, or 37.5%, of patients with pathogenic variants in this study) does not preclude an X-linked disorder. In routine diagnostics, X-inactivation studies can thus be helpful in directing the order of genetic tests, but they should be applied with caution.

The XLID genes *KDM5C* and *SMC1A* had been reported to escape X-inactivation.^{26,27} As in the girl with the *IQSEC2* mutation (patient 19) mentioned above, the mechanisms that caused skewing in female members of families 8 and 9 (with *KDM5C* mutations) and in family 17 (with mutation in *SMC1A*) await their elucidation.

In conclusion, targeted NGS-based analysis is a powerful tool to detect disease-causing mutations in XLID patients. The results of this study broaden the mutational spectrum of rare XLID genes, and they provide insight into the relative contribution of individual genes in this highly heterogeneous disorder. The application of this approach to large patient cohorts in routine diagnostics will provide the basis for a comprehensive and detailed picture of the genetic landscape of XLID.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the patients and their parents for their support. We also thank Juliane Gohlke for excellent technical assistance. This work was supported in part by EU funding (TECHGENE, FP7-Health 2007, B223143) and the German Intellectual Disability Network through a grant from the German Ministry of Education and Research (01GS08160).

- Ellison J, Rosenfeld J, Shaffer L: Genetic basis of intellectual disability. *Annu Rev Med* 2013; **64**: 441–450.
- 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.
- Lubs H, Stevenson R, Schwartz C: Fragile X and X-linked intellectual disability: four decades of discovery. *Am J Hum Genet* 2012; **90**: 579–590.
- Plenge R, Stevenson R, Lubs H, Schwartz C, Willard H: Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders. *Am J Hum Genet* 2002; **71**: 168–173.
- Piton A, Redin C, Mandel J: 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.
- Ropers H: Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 2010; **11**: 161–187.
- Rehm H: Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 2013; **14**: 295–300.
- Martin M: Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 2011; **17**: 10–12.
- Li H, Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; **26**: 589–595.
- Lunter G, Goodson M: Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res* 2011; **21**: 936–939.
- Li H: A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 2011; **27**: 2987–2993.
- Wang K, Li M, Hakonarson H: ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; **38**: e164.
- Saitou H, Nishimura T, Muramatsu K *et al*: De novo mutations in the autophagy gene *WDR45* cause static encephalopathy of childhood with neurodegeneration in adulthood. *Nat Genet* 2013; **45**: 445–449.
- Haack T, Hogarth P, Krueger M *et al*: Exome sequencing reveals de novo *WDR45* mutations causing a phenotypically distinct, X-linked dominant form of NBIA. *Am J Hum Genet* 2012; **91**: 1144–1149.
- Tarpey P, 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.
- Rinaldi C, Grunseich C, Sevrioukova I *et al*: Cowchock syndrome is associated with a mutation in apoptosis-inducing factor. *Am J Hum Genet* 2012; **91**: 1095–1102.
- van der Maarel S, Scholten I, Huber I *et al*: Cloning and characterization of *DXS6673E*, a candidate gene for X-linked mental retardation in Xq13.1. *Hum Mol Genet* 1996; **5**: 887–897.
- Lederer D, Grisart B, Digilio M *et al*: Deletion of *KDM6A*, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. *Am J Hum Genet* 2012; **90**: 119–124.
- Plon S, Eccles D, Easton D *et al*: Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 2008; **29**: 1282–1291.
- Allen R, Zoghbi H, Moseley A, Rosenblatt H, Belmont J: Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 1992; **51**: 1229–1239.
- Whibley A, Plagnol V, Tarpey P *et al*: Fine-scale survey of X chromosome copy number variants and indels underlying intellectual disability. *Am J Hum Genet* 2010; **87**: 173–188.
- Riess A, Rossier E, Krüger R *et al*: Novel *SLC9A6* mutations in two families with Christianson syndrome. *Clin Genet* 2013; **83**: 596–597.
- Shoubridge C, Tarpey P, Abidi F *et al*: Mutations in the guanine nucleotide exchange factor gene *IQSEC2* cause nonsyndromic intellectual disability. *Nat Genet* 2010; **42**: 486–488.
- Tran Mau-Them F, Willems M, Albrecht B *et al*: Expanding the phenotype of *IQSEC2* mutations: truncating mutations in severe intellectual disability. *Eur J Hum Genet* 2014; **22**: 289–292.
- Epi4K C, Allen A, Berkovic S *et al*: De novo mutations in epileptic encephalopathies. *Nature* 2013; **501**: 217–221.
- Brown CJ, Miller AP, Carrel L, Rupert JL, Davies KE, Willard HF: The *DXS423E* gene in Xp11.21 escapes X chromosome inactivation. *Hum Mol Genet* 1995; **4**: 251–255.
- Li N, Carrel L: Escape from X chromosome inactivation is an intrinsic property of the *Jarid1c* locus. *Proc Natl Acad Sci USA* 2008; **105**: 17055–17060.

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