Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing

Claire Redin, Bénédicte Gérard, Julia Lauer, Yvan Herenger, Jean Muller, Angélique Quartier, Alice Masurel-Paulet, Marjolaine Willems, Gaétan Lesca, Salima El-Chehadeh, Stéphanie Le Gras, Serge Vicaire, Muriel Philipp, Michaël Dumas, Véronique Geoffroy, Claire Feger, Nicolas Haumesser, Yves Alembik, Magalie Barth, Dominique Bonneau, Estelle Colin, Hélène Dollfus, Bérénice Doray, Marie-Ange Delrue, Valérie Drouin-Garraud, Elisabeth Flori, Mélanie Fradin, Christine Francannet, Alice Goldemberg, Serge Lombroso, Michèle Mathieu-Dramard, Dominique Martin-Coignon, Didier Lacombe, Gilles Morin, Anne Polge, Sylvie Sukno, Christel Thauvin-Robinet, Julien Thevenon, Martine Doco-Fenzy, David Genevieve, Pierre Sarda, Patrick Edey, Bertrand Isidor, Bernard Jost, Laurence Olivier-Faivre, Jean-Louis Mandel, Amélie Piton

ABSTRACT

Background Intellectual disability (ID) is characterised by an extreme genetic heterogeneity. Several hundred genes have been associated to monogenic forms of ID, considerably complicating molecular diagnostics. Trio-exome sequencing was recently proposed as a diagnostic approach, yet remains costly for a general implementation.

Methods We report the alternative strategy of targeted high-throughput sequencing of 217 genes in which mutations had been reported in patients with ID or autism as the major clinical concern. We analysed 106 patients with ID of unknown aetiology following array-CGH analysis and other genetic investigations. Ninety per cent of these patients were males, and 75% of them had undergone previous candidate gene or CGH analyses and other genetic investigations.

Results We identified 26 causative mutations: 16 in X-linked genes (ATRX, CUL4B, DMD, FMR1, HCF1, IL1RAPL1, IQSEC2, KDM5C, MAOA, MECP2, SLC9A6, SLC16A2, PHF8) and 10 de novo in autosomal-dominant genes (DYSR1K1A, GRIN1, MED13L, TCF4, RAII, SHANK3, SLC2A1, SYNGAP1). We also detected four possibly causative mutations (eg, in NLGN3) requiring further investigations. We present detailed reasoning for assigning causality for each mutation, and associated patients’ clinical information. Some genes were hit more than once in our cohort, suggesting they correspond to more frequent ID-associated conditions (KDM5C, MECP2, DYSR1K1A, TCF4). We highlight some unexpected genotype to phenotype correlations, with causative mutations being identified in genes associated to defined syndromes in patients deviating from the classic phenotype (DMD, TCF4, MECP2). We also bring additional supportive (HCF1, MED13L) or unsupportive (SHROOM4, SRY2) evidences for the implication of previous candidate genes or mutations in cognitive disorders.

Conclusions With a diagnostic yield of 25% targeted sequencing appears relevant as a first intention test for the diagnosis of ID, but importantly will also contribute to a better understanding regarding the specific contribution of the many genes implicated in ID and autism.

INTRODUCTION

Intellectual disability (ID) is a common neurodevelopmental disorder reported in 1.5–2% of children and adolescents. ID is defined by significant limitations in both intellectual functioning and adaptive behaviour with onset before the age of 18. Different classes of ID are conventionally defined upon IQ values (severe or profound, <50; moderate, 25–49; mild, 50–70). However, in routine genetic practice, clinical assessment mainly based on records of developmental history, speech acquisition and patients’ autonomy is used for classification in such subcategories.

Causes of ID can be environmental, genetic or multifactorial. Single genetic events are thought to account for a majority of cases, varying from large chromosomal anomalies or copy number variants (CNVs) affecting several genes to point mutations in single genes. These latter monogenic forms are characterised by an extreme genetic heterogeneity, with a hundred genes described as implicated in X-linked ID (XLD), and more associated to autosomal-recessive or autosomal-dominant forms. Altogether there are more than 500 genes proposed to cause ID with high penetrance when mutated underlying a phenotypic heterogeneity of the same extent in both severity and associated symptoms. This genetic heterogeneity has long limited the diagnostic offer for patients and families, which
was often restricted to fragile-X (MIM 300624) testing, array-CGH (comparative genomic hybridization) analysis and generic metabolic tests (see online supplementary figure S1). It may be complemented by sequencing a few genes associated to a specific syndrome evoked by patients’ phenotype, yet the diagnostic yield remains low (1–2% for the recurrent fragile-X mutation; 10–15% for array-CGH and chromosomal analyses, higher in highly syndromic patients). A majority of patients remain therefore without molecular diagnosis, while it is of crucial importance for establishing recurrence risks and providing genetic counselling in the family. Moreover, such diagnosis often has direct consequences for the medical prognosis of patients or their optimised healthcare, and even (yet in still a minority of cases) can indicate specific therapeutic options.

To obviate this low diagnostic yield, we developed the simultaneous targeted sequencing of protein-coding exons of 217 genes associated with ID or autism spectrum disorders (ASDs) as primary clinically significant feature: 99 located on the X-chromosome, 118 on the autosomes. We report here the results of such strategy on a cohort of 106 ID patients with or without associated autistic-like features, negative for array-CGH, fragile-X and other specific genetic analyses. A causal mutation was detected in 25% of these patients, regardless the severity of their cognitive impairment. We illustrate cases in which the molecular diagnosis was immediately established, as well as other more complex situations. This highlights the challenge of interpreting variants generated by NGS technologies, already from targeted approaches restricted to a few hundred genes. This work demonstrates that a targeted sequencing approach is highly efficient for the diagnosis of ID, but also allows refining the clinical spectrum associated with mutations in certain genes, and confirming or questioning the involvement of other genes in cognitive disorders.

METHODS
Cohort of patients
DNA samples from 106 patients were addressed for testing through clinical geneticists from 16 public hospitals in France. Inclusion criteria for the patients were to be negative for the recurrent fragile-X mutation and for pathogenic CNVs via array-CGH testing, and availability of DNA samples from family members in order to conclude upon molecular findings. Additional specific genetic investigations had been performed on a majority of patients (on average: two genes tested per patient, table 1). Presence of multiple major congenital anomalies or suspected mitochondrial/peroxisomal disorders was an exclusion criterion. Clinical data were recorded following a standardised clinical record highlighting prenatal history, developmental history, neurological and behavioural disorders. ID severity was assessed by medical geneticists upon clinical evaluation and was not a discriminating inclusion criterion, although we encouraged inclusion of probands with moderate or severe ID. This study was approved by the local Ethics Committee of the Strasbourg University Hospital (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale (CCPPRB). For all patients, a written informed consent for genetic testing was obtained from their legal representative.

Targeted genes and capture design
The 217 selected genes include 99 genes associated to XLID and 118 genes located on autosomes, implicated in dominant (45), recessive (66) or complex (7) forms of ID (see online supplementary table S8, further justification on gene selection is given in online supplementary methods). We targeted all protein-coding exons of these genes, including 20 bp of intronic flanking sequences. The overall size of targeted regions is 1.034 Mbp. Corresponding 120 bp RNA baits were designed using SureDesign (https://earray.chem.agilent.com/suredesign/).

Library preparation and sequencing
DNA samples were extracted from peripheral blood or saliva. Sequencing libraries were prepared as described previously, performing individual in-solution SureSelect capture reaction for each DNA sample (Agilent, Santa Clara, California, USA). Paired-end sequencing (2×101-bp) was performed on an Illumina HiSeq 2000/2500, multiplexing up to 16 samples per sequencing lane.

Bioinformatic pipeline and variant ranking
Read mapping, variant calling and annotation was performed as described previously. Detected variants (short indels and single nucleotide variants (SNVs)) were ranked by VaRank (an in-house developed script), which incorporates the annotations retrieved by alamut-HT (putative effect on the protein, conservation scores, splice site predictions, allelic frequency in the 106 patients and in control cohorts such as Exome Variant Server (EVS) or 1000 genomes). Candidate variants were selected when harbouring a frequency compatible with the incidence of the disease: expectedly accounting for less than 0.1% of all ID cases, aka resulting in a disease frequency <0.002%. Using the EVS population as a subset of the general population, candidate variants were thus retrieved when reported in EVS: with a minor allele frequency lower than 0.45% for variants in autosomal-recessive genes (ie, with a frequency of homozygotes <0.002%), in no more than one carrier for variants in autosomal-dominant genes or in no more than one male for variants in X-linked genes (as we cannot exclude, although it is unlikely, that a particular carrier from the general population might have mild ID). We concomitantly excluded variants present more than twice in the cohort of 106 patients. Remaining variants predicted as potentially pathogenic and fitting with the mode of inheritance associated to the affected gene were further tested for validation (see online supplementary table S9).

CNVs detection pipeline
Putative heterozygous/homozygous/hemizygous structural variants or CNVs were highlighted using the previously described method based on a depth-of-coverage comparison between the index sample and eight other random samples from the same sequencing lane. For the X-chromosome, coverage was normalised according to the number of X-chromosomes of the patient.

Mutation validation
All candidate mutations were validated by Sanger sequencing and co-segregation analyses were performed as extensively as possible. Putative splicing mutations were confirmed either using a minigene in vitro assay with the SPL3B plasmid as described previously or using patients’ fibroblasts or blood RNA when available. For apparent de novo variants, pedigree concordance was checked using polymorphic microsatellite markers (PowerPlex 16HS System, Promega). Mutations were considered as certainly causative when no doubt remained regarding their pathogenicity and an unambiguous diagnosis could thus be established. Such mutations co-segregated with the disease status in the family and were either truncating mutations or missense mutations that had been previously
convincingly published or that we confirmed with functional analyses. Mutations were considered as potentially causative when they appeared to co-segregate with ID in available members of the family and were predicted to be damaging, but further functional studies are needed to prove unambiguously their pathogenicity.

RESULTS

High-quality sequencing data ensure low rates of false-positive/negative calls for SNVs, indels and CNVs

Our strategy allowed generating a high-quality sequencing dataset, with a mean depth of coverage of 350× and an average per patient of 97.7% of targeted regions being well covered (>40×; see online supplementary table S1). Such coverage ensures a sensitivity of 99–99.9% of detecting SNVs and indels at any allelic state (Illumina Technical Note). We further assessed the sensitivity of SNV detection by comparing allelic states of SNPs detected by SNP-array (Affymetrix SNP Array 6.0) with the corresponding sites located in the targeted sequencing data of two patients and found no false-negatives after Sanger sequencing validation of six SNPs showing discrepant allelic states between both methods (452 SNPs analysed). Interestingly, for these six SNPs, Sanger sequencing results were always in favour of targeted sequencing data suggesting a much higher accuracy (data not shown). No false positive was detected out of the 80 candidate variants located in well-covered regions that were tested for confirmation by Sanger sequencing.

This high-sequencing depth also ensured reliable CNVs calling. All CNVs detected by our pipeline were validated by Sanger sequencing, qPCR and/or confirmed retrospectively when looking at array-CGH data (see online supplementary table S2). Some were not initially mentioned in the array-CGH report because they were covered by only a few SNP probes (≤2 deleted probes, under the detection threshold).

Very few regions (a total of 3.9 kb, only 1.8 kb being protein coding) appear consistently poorly covered (coverage <40× in >90% of the samples; see online supplementary table S3). Those are mainly first exons or highly GC-rich regions that are a well-known burden in such capture strategy.

Table 1  Description of the cohort of 106 patients with intellectual disability (ID) and global diagnostic results

|                | Cohort (n=106) | With conclusive genetic diagnosis (n=26) | Yield (per category) |
|----------------|---------------|----------------------------------------|---------------------|
| Gender         |               | XLID | ADID | Total |                      |
| Male           | 96 (91%)      | 14   | 8    | 22    | 22/96 (23%)          |
| Female         | 10 (9%)       | 2    | 2    | 4     | 4/10 (40%)           |
| Total          | 106           | 16   | 10   | 26    | 26/106 (25%)         |
| Age            |               |      |      |       |                      |
| [0–10]         | 57 (54%)      | 9    | 5    | 14    | 14/57 (25%)          |
| [10–20]        | 31 (29%)      | 5    | 4    | 9     | 9/31 (29%)           |
| >20            | 18 (17%)      | 2    | 1    | 3     | 3/18 (17%)           |
| Sporadic cases *|              |      |      |       |                      |
| Female         | 8 (7%)        | 2    | 2    | 4     | 4/8 (50%)            |
| Male           | 72 (68%)      | 7    | 7    | 14    | 14/72 (19%)          |
| Total          | 80 (75%)      | 9    | 9    | 18    | 18/80 (23%)          |
| Familial history|             |      |      |       |                      |
| Male sib-pairs | 12            | 5    | 0    | 5     | 5/12 (42%)           |
| Possible XLID† | 8             | 2    | 1    | 3     | 3/8 (38%)            |
| Other (non-X-linked) | 6 | 0    | 0    | 0     | 0/6 (0%)            |
| Total          | 26 (25%)      | 7    | 1    | 8     | 8/26 (31%)           |
| Consanguinity  | 3 (3%)        | 0    | 0    | 0     | 0/3 (0%)            |
| ID severity    |               |      |      |       |                      |
| Mild/borderline| 12 (11%)      | 3    | 0    | 3     | 3/12 (25%)           |
| Moderate       | 49 (46%)      | 4    | 6    | 10    | 10/49 (20%)          |
| Severe         | 45 (42%)      | 9    | 4    | 13    | 13/45 (29%)          |
| Comorbidity    |               |      |      |       |                      |
| Microcephaly (<−2 SD) | 14 (13%) | 3    | 2    | 5     | 5/14 (36%)           |
| Epilepsy       | 28 (26%)      | 5    | 2    | 6     | 6/28 (21%)           |
| Autistic traits| 34 (32%)      | 7    | 2    | 9     | 9/34 (26%)           |
| Hypotonia      | 36 (34%)      | 4    | 4    | 8     | 8/36 (22%)           |
| Previous exploration |   |      |      |       |                      |
| CGH            | 106 (100%)    |      |      |       |                      |
| Fragile-X      | 105 (99%)     |      |      |       |                      |
| Karyotype      | 99 (96%)      |      |      |       |                      |
| # Other genetic tests (mean per patient) | 2 | | | | |
| MRI            | 53 (50%)      |      |      |       |                      |
| Metabolism‡    | 82 (77%)      |      |      |       |                      |

* No familial first degree ID.
† Affected male relatives.
‡ At least one biochemical test performed.
XLID, X-linked ID; in italic: total number per category.

Redin C, et al. J Med Genet 2014;51:724–736. doi:10.1136/jmedgenet-2014-102554
Cohort description and diagnostic yield

Patients harboured various degrees of cognitive impairment, although with a higher proportion of moderate or severe forms (46% and 42% respectively; table 1). The cohort was highly enriched in males. Among male probands, 68% were sporadic cases, the remaining had familial history of cognitive impairment mainly evocative of an X-linked mode of transmission (table 1). We detected certainly causative mutations in 26/106 patients (table 2; see online supplementary figures S2–S19), leading to an overall diagnostic yield of 25% for the entire cohort (from 23% for sporadic cases to 31% for familial cases). Unexpectedly, the diagnostic yield appears unrelated to the severity of ID in patients (table 1).

Table 2  List of all causative/possibly causative mutations identified in our cohort

| Patient ID | Sex | Gene | Mutation | Inheritance | Mode of inheritance | Degree of ID | Consistency with classic phenotype | See online supplementary figure # |
|------------|-----|------|----------|-------------|--------------------|--------------|-------------------------------|-----------------------------------|
| APN-58     | M   | Dyrk1a | chr21:g.38858865C>T; c.613C>T; p.Arg205*; htz | De novo | AD | ++ | Yes | S2 |
| APN-87     | M   | Dyrk1a | chr21:g.38858873del; c.621_624delinsGAA; p.Glu204Asufs*3; htz | De novo | AD | ++ | Yes | S2 |
| APN-63     | M   | Grin1 | chr9:g.140056661C>G; c.1733C>G; p.Val578Arg; htz | De novo | AD | +++ | Yes | S3 |
| APN-14     | M   | Med13l | chr12:g.116406845_116406852del; c.6118_6126del; p.Gly2040Asufs*32; htz | De novo | AD | ++ | Partially | S4 |
| APN-46     | M   | RaI1 | chr17:g.17698594_17698598del; c.2332_2336del; p.Gly778Glnfs*7; htz | De novo | AD | ++ | Partially | S5 |
| APN-122    | F   | Shank3 | chr1:g.51159168_51159183dup; c.520C>T; p.Arg174*; htz | De novo | AD | ++ | No | S9 |
| APN-38     | M   | Slc2a1 | chr12:g.116406845_116406852del; c.6118_6126del; p.Gly2040Asufs*32; htz | De novo | AD | ++ | Partially | S4 |
| APN-139    | M   | Syngap1 | chr18:g.3017622; p.Glu242*; htz | Inherited (Ma) | XL | ++ | No | S9 |
| APN-41     | M   | TCF4 | chr18:g.3017622; p.Glu242*; htz | Inherited (Ma) | XL | ++ | No | S9 |
| APN-117    | F   | TCF4 | chr18:g.3017622; p.Glu242*; htz | Inherited (Ma) | XL | ++ | No | S9 |
| APN-138    | M   | Atrx | chrX: g.76972632G>G; c.109C>T; p.Arg37* (rs12245108); hzm | Inherited (Ma) | XL | ++ | Yes | S10 |
| APN-137    | M   | CUL4B | chrX: g.119861009_119861010del; c.811_812del; p.Gln271Aspfs*11; hzm | Inherited (Ma) | XL | ++ | Partially | S11 |
| APN-42     | M   | Dmd | chrX: g.31017622; p.Glu242*; htz | De novo | AD | ++ | Partially | S12 |
| APN-26     | M   | Fmr1 | chrX: g.31017622; p.Glu242*; htz | De novo | AD | ++ | Partially | S13 |
| APN-113    | M   | Hcfc1; (Atrx) | chrX: g.76973735G>G; c.1013C>G; p.Ser338Cys; hzm | Inherited (Ma) | XL | ++ | Partially | S14 |
| APN-82     | M   | Il1rapl1 | chrX: g.29935696_29935705del; c.894_903del; p.Trp299Thrfs*18; hzm | Inherited (Ma) | XL | ++ | Yes | S14 |
| APN-68     | M   | IQSEC2 | chrX: g.23288395G>G; c.3097C>T; p.Gln1033*; hzm | De novo | XL | ++ | Yes | S15 |
| APN-34     | M   | Kdm5c | chrX: g.53240784dup; c.1296dup; p.Glu433*; hzm | Inherited (Ma) | XL | ++ | Partially | S16 |
| APN-16     | M   | Moca | chrX: g.34309092; g.3430935delinsT; c.797_798delinsT; p.Gly266Phe; hzm | Inherited (Ma) | XL | ++ | Partially | S17 |
| APN-130    | F   | Mecp2 | chrX: g.15329636G>G; c.916C>T; p.Arg306Cys (rs28935468); hzm | Inherited (Ma) | XL | ++ | Partially | S18 |
| APN-142    | F   | Mecp2 | chrX: g.1532967777G>G; c.502C>T; p.Arg168* (rs81748421); hzm | Inherited (Ma) | XL | ++ | Partially | S19 |
| APN-3      | M   | Mecp2 | chrX: g.15329636G>G; c.916C>T; p.Arg306Cys (rs28935468); hzm | Inherited (Ma) | XL | ++ | Partially | S17 |
| APN-105    | M   | Phfr; (DOCR8) | chrX: g.40028538G>G; c.1249+5G>C; p.Tyr406Phefs*24; hzm | Inherited (Ma); (de novo) | XL | ++ | Partially | S18 |
| APN-43     | M   | Slc9a6 | chrX: g.135080258; 135080262del; c.526_9,526_5del; p; splice disrupted; hzm | Inherited (Ma) | XL | ++ | Yes | S19 |
| APN-110    | M   | Slc16a2 | chrX: g.73749677C>T; c.1412T>C; p.Leu471Pro (rs122455132); hzm | Inherited (Ma) | XL | ++ | Partially | S19 |

Continued
Cognitive and behavioural genetics

### Table 2 Continued

| Patient ID | Sex | Gene  | Mutation                                                                 | Inheritance                  | Mode of inheritance | Degree of ID | Consistency with classic phenotype | See online supplementary figure # |
|------------|-----|-------|--------------------------------------------------------------------------|------------------------------|---------------------|-------------|-----------------------------------|-----------------------------------|
| APN-131    | M   | SLC2A1| chr1:g.43392779de; c.1412delG; p.Gly471GluX4; 37; hz; (chr16:g.89348867G>T; c.4083C>A; p.His1361Gln; hz) | Inherited (Pa);              | AD                  | +++         | Partially                        | 57                                |
| APN-101    | M   | TCF4  | chr18:g.52899070C>T> A; p.Arg495_Gly496insAla?; hz                       | De novo ? †                  | AD                  | ++          | No                               | 59                                |
| APN-99     | M   | NLGN3 | chrXq.70389249C>T> A; c.1849C>T; p.Arg617Trp; hemz                     | Inherited (Ma)               | XL                  | +++         | Yes                              | S20                               |
| APN-70     | M   | PQB1  | chrXq.48762094C>T> A; c.731C>T; p.Pro244Leu; hemz                     | Inherited (Ma)               | XL                  | ++          | No                               | S21                                |

*Present in the three brothers. Mother untested, but most probably maternally inherited.
†Absent from the mother, deceased father (untested).

In bold: mutations previously reported in other patients.

−: no ID; +: mild ID; ++: moderate ID; +++: severe ID.

AD, autosomal dominant; F, female; hzm, hemizygous; hzt, heterozygous; ID, intellectual disability; M, male; Ma, maternally-inherited; MAOA, Monoamine Oxidase A enzyme; Pa, paternally inherited; XL, X-linked.

Sixteen mutations are located in genes of the X-chromosome: 14 point mutations or small indels (in ATRX, CUL4B, DMD, HCFC1, IL1RAPL1, IQSEC2, KDM5C, XMECP2, MAOA, PHF8, SLC9A6, SLC16A2), as well as two larger pathogenic events (one hemizygous complex rearrangement in MECP2, one hemizygous exon deletion in FMR1; see online supplementary figures S17 and S13). We identified 10 de novo point mutations or small indels in genes involved in autosomal-dominant/haploinsufficient forms of ID (in DYSK41A2, GRIN1, MED13L, RA11, SHANK3, SYNGAP1, SLC2A1, TCF4x2). In four other patients, we identified potentially causative mutations (in NLGN3, PQB1, SLC2A1 and TCF4; see online supplementary figures S20, S21, S7 and S9, respectively), whose implication in cognitive impairment has to be further confirmed. Finally, missense variants that appeared at first likely to be pathogenic, notably based on the very high evolutionary conservation of the affected residues or on previous publications, appeared excluded as causal after further segregation analysis (in FLNA, FMR1, HUWE1 or MECP2, see online supplementary figure S22).

**Atypical type of mutations**

Among the 26 certainly causative mutations identified, some were surprising by the nature of the mutation itself. In a boy with severe encephalopathy, epilepsy, hypotonia and microcephaly, we detected a highly complex rearrangement in exon #4 of MECP2, involving a 139 bp deletion flanked by the insertion of two sequences in inverted orientation derived from intron #2 still keeping the reading frame downstream of the rearrangement. Such event is inherited from the proband’s mother who presents with speech delay and dyslexia (see online supplementary figure S17). This exon is known to be the one accumulating most mutations in patients, especially the 3’ half, which is a recombination hotspot and the target of several deletions/duplications and less frequently inversions. 16–18

We report here for the first time an intragenic deletion affecting FMR1 outside of the promotor/exon #1 region that is the target of the fragile-X syndrome CGG-expansion. Very few point mutations have been reported in coding regions. We identified a complete deletion of the last exon of FMR1 in one patient and his two affected brothers unevenly presenting with clinical features of fragile-X syndrome (see online supplementary figure S13).

We also identified a patient carrying a maternally inherited 10 bp deletion causing a frameshift in exon #7 of IL1RAPL1, while unexpectedly his affected brother bears a de novo deletion of the full exon highlighting that affected relatives may carry distinct mutations. We propose that small 10 bp deletion may have created a sequence conformation favouring further instability and leading to the larger deletion observed in the second brother (see online supplementary figure S14). Indeed, a large proportion of IL1RAPL1 causative mutations are intragenic exon deletions or pericentric inversions,19 supporting that IL1RAPL1 region is highly susceptible to recombination events.

Lastly, we identified a de novo Pro578Arg missense mutation in GRIN1 in a male proband with severe ID, hypotonia, feeding disorders and very poor speech but no epilepsy, a phenotype similar to that associated to GRIN1 missense mutations in some patients. 20 21 A maternal uncle presented with similar features except for the poor speech and hypotonia (see online supplementary figure S3), initially suggesting an associated X-linked mode of inheritance. The finding of a de novo deleterious missense mutation excludes this latter hypothesis, further highlighting the prevalence of phenocopies in cognitive disorders.

**Genotype–phenotype correlations: from expected to unexpected**

The majority of certainly causative mutations were identified in patients whose clinical phenotype was retrospectively consistent with previous reports (table 2). For instance, the proband carrying a truncating mutation in IQSEC2 presents with severe ID, no speech, motor developmental delay, severe epilepsy, strabismus and autistic features (see online supplementary figure S15), which matches the recently proposed clinical spectrum associated to mutations in this gene. 22 23 Likewise, DYSK1A was originally found disrupted by translocations or deleted in several patients with ID and microcephaly. 24–26 More recently, truncating mutations in this gene were shown to cause ID associated with primary microcephaly (sometimes borderline at −2 SD), growth retardation, developmental delay, facial dysmorphic traits, seizures and major feeding difficulties, with or without

---

Redin C, et al. J Med Genet 2014;51:724–736. doi:10.1136/jmedgenet-2014-102554
associated autism.27–31 We report here two novel de novo truncating mutations in patients with similar clinical features but no epilepsy (see online supplementary figure S2). Nonetheless, in a few other cases (eg, mutations in RAII or MECP2), the probands lacked some clinical features, thus the corresponding diagnosis of Smith-Magenis (MIM 182290) or Rett (MIM 312750) syndrome was not evoked by experienced clinical geneticists (see online supplementary figures S5 and S17). For instance, among the three patients with MECP2 mutations, one female had the classic Rett phenotype, the other female proband presented with non-classic Rett phenotype (no regression episode, no hand-flapping), while the rarity of reports of MECP2 encephalopathies in adolescent males precluded suspicion of the involvement of that gene in a patient (APN-3) who had undergone rather extensive prior genetic testing.

A few detected mutations were unexpected as they were detected in patients whose phenotype did not match previous descriptions. For instance, mutations in TCF4 mainly cause Pitt-Hopkins syndrome (PHS, MIM 610954) characterised by severe motor retardation, absence of speech, characteristic dysmorphic traits, autist features, intestinal problems and hyperventilation.32 We here describe truncating TCF4 mutations in two patients, one with clinical manifestations highly suggestive of PHS, the other with less syndromic manifestations and no dysmorphic traits (see online supplementary figure S9). TCF4 mutations were already reported in patients with non-syndromic ID, suggesting that such mutations were likely to be underdiagnosed.33

Another patient and his affected brother both carry a distal frameshift mutation in DMD affecting the major muscle transcript encoding the dystrophin protein associated to Duchenne or Becker muscular dystrophy (DMD, MIM 310200; BMD, MIM 300376), and the brain-specific isoform Dp71. The index case presents with moderate ID, psychomotor retardation, no speech, behavioural disorders, dysmorphic traits but strikingly no muscular phenotype (see online supplementary figure S12). His brother harbours a milder phenotype with additional hypotonia and cerebellar dysplasia. Both harbour borderline-high CPK levels. The association of cognitive impairment with DMD/BMD has been extensively reported and correlated to CPK levels. The association of cognitive impairment with tonia and cerebellar dysplasia. Both harbour borderline-high His brother harbours a milder phenotype with additional hypo- no muscular phenotype (see online supplementary speech, behavioural disorders, dysmorphic traits but strikingly case presents with moderate ID, psychomotor retardation, no muscular phenotype (see online supplementary

Confirmation of candidate genes for cognitive disorders
Some selected genes were only candidate ID or ASD genes at the time of the design, with single pieces of evidence in the literature. The identification of additional mutations in patients with similar phenotype definitively confirms their implication in cognitive disorders.

We reported a damaging missense affecting the function of the Monoamine Oxidase A enzyme (MAOA),31 which implicated for the first time in 20 years the implication of MAOA in autism/ID associated to significant behavioural disorders.31,32 We also identified a probably pathogenic missense variant in NLGN3 in a male and his cousin, both presenting with ID and autism (see online supplementary figure S20). In silico predictions, high conservation of the mutated residue across all neuro- ligin paralogs, and familial analysis are altogether in favour of a pathogenicity of this missense change. A definitive functional effect of this missense still needs to be demonstrated to clearly establish the diagnosis. The implication of this gene was never replicated since the initial publication.33 Although screened in several cohorts with comparable phenotypes.34–36

Similarly, we identified a novel truncating point mutation in MED13L confirming the implication of this gene in ID (see online supplementary figure S4). Disruption of MED13L was initially associated with transposition of the great arteries (TGA), associated to ID in a single case with a chromosomal translocation.34 A homozygous missense mutation was then identified in two siblings from a consanguineous family presenting with non-syndromic ID, suggesting an implication of the gene in autosomal-recessive forms of ID justifying its selection in our panel.34 A total of five patients were more recently described with de novo intragenic CNVs or point mutations affecting MED13L, delineating a recognisable MED13L-haploinsufficiency syndrome characterised by hypotonia, moderate ID, variable cardiac defects, facial hypotonia and dysmorphic traits.35,36 Our patient with the MED13L mutation presents with concordant phenotype such as an open mouth appearance and muscular hypotonia but no cardiac defects. Due to the initially proposed autosomal-recessive mode of inheritance associated to ID, at first we did not consider this heterozygous truncating mutation as causative, as it may also have been the case for a heterozygous splicing mutation identified earlier in a large-scale exome sequencing study in one male with ASD.37 Altogether, these findings suggest that ID associated to MED13L-haploinsufficiency syndrome is a relatively frequent condition.35,36

Ambiguous mode of inheritance in ID-associated genes: the example of DEAF1
As for MED13L, the mode of inheritance associated to some genes is ambiguously described in literature. We identified a heterozygous variant affecting splicing in DEAF1 inherited from the asymptomatic mother in a patient presenting with severe ID, developmental delay, poor speech, pain resistance, dysmorphic features and aggressive behaviour (figure 1), while the gene had been proposed as associated to autosomal-dominant forms of ID.7,8 The recent report of two additional individuals carrying de novo missense mutations narrowed the associated phenotype to moderate/severe ID, speech impairment, behavioural problems, high pain threshold, dysmorphic features and abnormal walking pattern, hence highly similar to the one of our proband.38 Although in vitro validation studies suggest that the reported missense variants lead to an impaired function of DEAF1, the authors concluded that they presumably act as dominant-negatives incapacitating both normal and mutant proteins since truncating variants had been observed in asymptomatic individuals.39 In parallel, a homozygous missense mutation clustering in the same SAND-domain with all three de novo missense mutations was reported in members of a consanguineous family presenting with ID, microcephaly and white matter abnormalities, therefore suggesting a possible autosomal-recessive mode of inheritance.40 The pathogenic mechanism associated to DEAF1 mutations is therefore unclear. Due to the highly similar clinical features of the herein reported proband and of probands carrying de novo missense mutations, the splice variant detected here may contribute to the phenotype of our patient, possibly through a recessive mode of inheritance (ie, acting in trans with another heterozygous variant) since haploinsufficiency appears tolerated in healthy individuals. Altogether those findings either suggest a similar phenotype for autosomal-dominant and autosomal-recessive mode of inheritance associated to DEAF1 mutations, or a universal autosomal-recessive mode of inheritance with a second variant that has not
Figure 1 Inherited disrupting splice variant in \textit{DEAF1}: what contribution to the phenotype? (A) Pedigree showing the maternally inherited splice variant in \textit{DEAF1} (c.290-3C>G); (B) prediction scores for the effect of the herein described variant on splicing (prediction scores for acceptor splice sites (ASS) as computed by SpliceSite Finder, MaxEntScan, NNSplice, GeneSplicer and Human Splicing Finder for the consensus ASS with either the wild-type or the mutated allele); (C) localisation of the variant, and its resulting effect on splicing in vitro (minigene construct): 90% of abnormal transcripts: 80% with entire exon #2 skipped, and 10% using the alternative ASS c.290-16 both leading to a frameshift and a premature stop codon.

Unsupportive evidences for proposed ID-associated mutations and genes

The identification in patients of a previously reported ID mutation should be considered with caution: indeed, as already discussed in previous publications, some ‘false-positive’ mutations are present among variants annotated as ‘pathogenic’ in dbSNP OMIM or in published reports.\textsuperscript{3,14} We here question the causative effect of a few previously proposed mutations (see online supplementary table S4). For instance, one missense variant identified in \textit{FLNA} (c.3872C>T, p.Pro1291Leu) was previously reported in a patient with FG syndrome (MIM 300321).\textsuperscript{57} This variant was identified here in a patient with a different phenotype (see online supplementary figure S22) and is also reported in one male in EVS (presumably not presenting with cognitive disorders), raising doubt about its pathogenicity.

The identification of non-segregating truncating variants (ie, detected both in patients and healthy relatives) can also challenge the implication of genes in X-linked and autosomal-dominant forms of ID. In one family, we identified a frameshift variant in \textit{SRPX2} in a male proband. It is most likely inherited from the deceased asymptomatic maternal grandfather since it was also detected in the mother and in three maternal aunts yet absent from the maternal grandmother. Despite recent functional evidences regarding the role of \textit{SRPX2} in brain development,\textsuperscript{58,59} its definitive implication in cognitive disorders has already been questioned following the presence of the initially proposed mutations in EVS\textsuperscript{53} and in control individuals,\textsuperscript{60} and subsequently to the identification of a missense mutation in \textit{GRIN2A} co-segregating with the epileptic status in the initial \textit{SRPX2} family.\textsuperscript{61} In another family, we identified a nonsense variant in \textit{SHROOM4} in a male proband yet also in his unaffected brothers, a finding that further challenges the implication of \textit{SHROOM4} in X-linked cognitive disorders (see online supplementary table S5, figure 2).\textsuperscript{14}

Patients carrying probably pathogenic variants in two ID-associated genes

In three unrelated patients, we identified candidate variants in two separate genes, requiring the evaluation of different scenarios: (a) one single contributor while the second variant is innocuous, (b) one major contributor while the second variant acts as a modifier, and (c) both variants are implicated in the phenotype and have a synergistic effect.

In one male proband, we identified a maternally inherited splice variant leading to a frameshift in the X-linked gene \textit{PHF8}, together with a de novo truncating variant in \textit{DOCK8} (see online supplementary figure S18). Considering both the weak evidences implicating \textit{DOCK8} in autosomal-dominant ID (two probands reported with translocations disrupting the gene), the clinical features consistent with a \textit{PHF8} mutation (even in the absence of cleft lip/palate) and the X-inactivation bias identified in the mother, the \textit{PHF8} variant alone is most likely to be responsible for the phenotype leaving the \textit{DOCK8} variant as probably innocuous.\textsuperscript{62,63}

In another male proband, two possibly causative missense variants were identified in the XLID genes \textit{ATRX} and \textit{HCFC1} (figure 3). Patients’ phenotype was not evocative of an \textit{ATRX} mutation (no dysmorphic traits, no urogenital abnormalities, absence of Heinz bodies), but perfectly matched the recent description of cobalamin-X metabolic disorder (\textit{cblX}, MIM 309541\textsuperscript{64}). We concluded to a causative effect of the \textit{HCFC1} mutation (being one of the \textit{cblX} recurrent mutations), but cannot exclude a contributory effect of the \textit{ATRX} variant based on current co-segregation data.

In the third family, the proband carries one distal truncating variant in \textit{SLC2A1} and one missense variant in \textit{ANKRD11}, each inherited from one asymptomatic parent. Some mutations in \textit{SLC2A1} have already been associated to GLUT-1 deficiency (MIM 606777/612126) with incomplete penetrance.\textsuperscript{65}
proband presents with evocative symptoms of both SLC2A1 and ANKR11 mutations (major hypotonia along with skeletal abnormalities) that might suggest a synergistic contribution of both variants to the phenotype (see online supplementary figure S7). No conclusion could be unambiguously drawn in this case, but such a di- or oligogenic mode of inheritance has already been proposed in neurodevelopmental disorders.66–68

**DISCUSSION**

Targeted sequencing of 217 genes in a cohort of 106 patients with unknown genetic aetiology of ID led to a conclusive diagnostic yield of 25%. A majority of causative mutations identified are located in XLID genes, which can be partially explained by our male-enriched cohort. When excluding familial cases, X-linked mutations are found in 7/72 sporadic male cases, matching the proposed figure that 10% of males with sporadic ID carry mutations in X-linked genes.69 Combining the results of recent trio-exome analyses leads to similar proportions considered (see online supplementary table S6), which may result in missing diagnostic result.92 02 8 70 71 If mutations in such ID genes are confirmed to account for a substantial number of patients, introducing as a diagnostic step massive multiplexed resequencing of such genes in large cohorts may be considered.

The extensive genetic and phenotypic heterogeneity of ID is the major hindrance for obtaining a precise molecular diagnosis. Direct sequencing strategies consisting in sequentially screening candidate genes are being replaced in diagnostic laboratories by more high-throughput NGS-based strategies: multiplex targeted sequencing of a few genes in large cohorts, selective targeted sequencing of up to several hundred genes, genome rearrangements and for CG-rich exons, remains in the research domain. Exome or full-genome strategies are more exhaustive alternatives and very attractive issues in the field of genetic diagnosis. Their universal approaches, whatever the clinical features, enhance the technical management of the workflow. However, the actual coverage proposed with the exome or full-genome strategies is still frequently insufficient (see online supplementary table S6), which may result in missing mutations.9 29 Also, the finding of a putative mutation in a gene never associated before to a given pathology marks the start of a research endeavour to validate the finding but is not per se a diagnostic result.

Targeted sequencing appears more appealing for well-defined pathologies in which most implicated genes have been uncovered or in clinically homogeneous entities (ie, Bardet-Biedl and related ciliopathies,13 retinal dystrophies,72 hearing loss,73 etc).
Figure 3  Patient carrying two probably damaging missense variants in HCFC1 and ATRX: one causative mutation and one modifier variant? (A) Family tree of patient APN-113: proband carries a missense mutation in HCFC1 (c.218C>T; p.Ala73Val) already reported in two patients with cblX as well as another missense variant in ATRX (c.1013C>G, p.Ser338Cys). Both variants are maternally inherited, absent from the unaffected brother, but carried by the younger brother who died of sudden death at 2 months old; (B) associated predictions for the recurrent pathogenic missense mutation in HCFC1 and the possible modifier in ATRX, showing putative pathogenicity and moderate nucleotide conservation; (C) representation of HCFC1 and its domains: kelch domains (K1–K5), Fn3 (fibronectin type 3), basic domain, HCF-proteolysis repeats (HCF-pro), acidic domain and nuclear localisation signal (NLS) domain. The initial mutations involved in milder non-syndromic ID are indicated above: a regulatory variant in the 5′UTR of HCFC1 was identified by targeted massive parallel resequencing in a family with probable X-linked ID (XLID) (MRX3), which had for long remained unsolved. This variant was disrupting the functional binding site of the transcription factor YY1 within the HCFC1 promoter region, leading to an upregulation of its expression in lymphoblastoid cells. Subsequent screening of additional unsolved families identified one single co-segregating missense variant (c.674G>A; p.Ser225Asn) in HCFC1. The phenotype of both patients was rather mild: non-syndromic mild to moderate ID. In the bottom are indicated the mutations recently described in cblX patients; (D) the ATRX missense variant is located close but outside of the hotspot for disease-causing missense mutations (in the zinc-finger binding domain, in green) reported in patients with ATRX mutations, in red: mutations reported independently in at least two patients.
We show here that it is also a powerful alternative for diagnostic purposes in ID (see online supplementary table S6). The positive diagnostic yield in our cohort of 106 patients is of 25% overall and 21% for sporadic cases, which is similar to the initial reports with the trio-exome strategy (32 highly likely causative mutations in a total of 151 patients, a 21% yield; see online supplementary table S7). Indeed, most of the mutations identified by exome sequencing affect genes included in our panel and would thus have been detected with our strategy. Our diagnostic yield is slightly lower than the 30% yield (on 211 patients) we now calculated from pooling the three published exome studies (including also29 and using the updated data for very recently reported in Gilissen et al), a difference that can be accounted for by the recently recognised ID genes missing in our panel (see online supplementary table S7). Our relatively high diagnostic yield is not due to an overestimation regarding the pathogenicity of the identified variants as we were highly stringent regarding the classification of certainly causative mutations (eg, we did not include all probably damaging missense variants located in XLID genes that were detected in males).

The high depth of coverage and the smaller portion of poorly covered regions achieved with our strategy ensure a high sensitivity and specificity of detecting pathogenic events in the regions of interest (SNVs, indels, but also CNVs, a topic that had not been addressed in previous exome studies). Also, the relative ease of variant analysis and smaller number of follow-up studies for candidate variants may contribute to this significant yield. As fewer candidate variants are identified per patient with this approach, they can also be analysed more thoroughly (putative effect on splicing for variants not affecting canonical splice sites, predictive impact on the protein through structural modeling, etc). With an even broader panel of genes (including, for instance, genes involved in ID associated to multiple congenital anomalies but found mutated in patients with a milder presentation), one can expect that the resulting diagnostic yield will be higher.

The limited number of sequenced genes with the targeted approach—restricted to those involved in cognitive disorders—also avoids the controversial issues raised by incidental findings. Nonetheless, the targeted sequencing approach will miss newly identified genes and theoretically prevents data reanalysis (ie, incorporating novel findings regarding genetics of ID). The major value of targeted sequencing for ID is that it should allow the application of such test to a much higher proportion of patients carrying mutations located in newly identified genes or, mutations in known genes yet that had been missed by previous bioinformatic pipelines.9 With an estimated cumulative diagnostic yield of 62% (although derived from a small cohort of 50 patients, for WGS analysis), whole-genome analyses of larger cohorts of patients may shed some more light regarding the contribution of the abovementioned hypotheses to explain the remaining proportion of patients without molecular diagnosis.

Author affiliations

1Département de Médecine translationnelle et Neurogénétique, IGBMC, CNRS UMR 7104/INSERM U964/Université de Strasbourg, Illkirch, France
2Chaire de Génétique Humaine, Collège de France, Illkirch, France
3Laboratoire de diagnostic génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France
4Centre de Génétique et Centre de Référence Anomalies du développement et Syndromes malformatifs, Hôpital d’Enfants, CHU de Dijon, Dijon, France
5Département de Génétique Médicale, Centre de Référence Maladies Rares Anomalies du Développement et Syndromes Malformatifs Sud-Languedoc Roussillon, Hôpital Arnaud de Villeneuve, Montpellier, France
6Département de Génétique Médicale, Hospices Civils de Lyon, Bron, France
7Plaftforme de Biopsies et Séquençage, IGBMC, CNRS UMR 7104/INSERM U964/Université de Strasbourg, Illkirch, France
8Plateforme de Bioinformatique de Strasbourg (BIPS), IGBMC, CNRS UMR 1074/INSERM U964/Université de Strasbourg, Illkirch, France
9Département de Génétique, CHU de Hautepierrre, Strasbourg, France
10Département de Biochimie et de Génétique, CHU d’Angers, Angers, France
11Laboratoire de Génétique Médicale, INSERM U1112, Faculté de Médecine de Strasbourg, Hôpitaux Universitaires de Strasbourg, Strasbourg, France
12CHU de Bordeaux, Génétique Médicale, Université de Bordeaux, Laboratoire MRGM, Bordeaux, France
13Département de Génétique Médicale, CHU de Rouen, Rouen, France
14Service de Génétique Médicale, Centre De Référence Anomalies du Développement, CHU de Rennes, Rennes, France
15Service de Génétique Médicale, Hôtel Dieu, Clermont-Ferrand, France
16Laboratoire de Biochimie, CHU de Nîmes, Nîmes, France
17Unité de Génétique Clinique, CHU d’Amiens, Amiens, France
18Service de Génétique, Centre Hospitalier, Le Mans, France
19Service de Neuropédiatrie, Hôpital Saint Vincent de Paul, Groupe Hospitalier de l’Institut Catholique Lillois, Faculté Libre de Médecine, Lille, France
20Service de Génétique, CHU de Reims, EA3801, Reims, France
21Service de Génétique Médicale, CHU de Nantes, Nantes, France

Acknowledgements: This study was supported by grants and fellowships from Fondation pour la Recherche Médicale, Agence de Biomedecine and Fondation Jerome Lejeune, APLM and CREGEMES. We thank the students who were involved in this project: Sébastien Kirsch, Audrey Creppy, Inès Bekkour and Grace Gan. We thank Nadège Calmels, Valérie Biancalana, Elisa Nourisson and all other members of the Genetic Diagnostic Laboratory of the Nouvel Hospital Civil (Strasbourg) for their help with patients’ DNA sample selection and preparation. We thank Cécile Pizot for the development of VaRank. We thank Damien Sanlaville, Christine Coubes, Delphine Héron, Sophie Naudin, James Lepinsaie and Marie-Line Bichon for their contribution to the recruitment of patients, and all medical interns or genetic counsellors who participated in this project. Finally, we warmly thank all patients and their families for their implication in this study.

Contributors: Study concept and design: CR, J-LM, AP. Clinical genetics investigations: AM-P, MW, GL, SE-C, YA, MB, DB, EC, HD, BD, M-AD, VD-G, EF, MF, CF, AG, SL, MM-D, DM-C, DL, GM, AP, SS, CT-R, JT, MD-F, DG, PS, PE, BI and LO-F. Acquisition, analysis and interpretation of data: CR, BG, JL, YH, JM, AQ, BI, CM and AP. Drafting of the manuscript: CR, J-LM and AP. Critical revision of the manuscript for important intellectual content: BG and JM. Obtained funding: J-LM and AP. Administrative, technical or material support: NH, MD, CF, VG, SLG, MP and SV. Study supervision: BG, J-LM, AP.

Competing interests: None.

Patient consent: Obtained.

Ethics approval: Comité de Protection des Personnes, France.

Provenance and peer review: Not commissioned; externally peer reviewed.
REFERENCES

1. Leonard H, Wen X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. Ment Retard Dev Disabil Res Rev. 2002;8:117–34.

2. Maulik PK, Macearenaus MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. Res Dev Disabil. 2011;32:1084–96.

3. de Ligt J, Willems MH, von Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Sifjouth A, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Visseres LE. Diagnose. The incidence of rearrangements in persons with severe intellectual disability. N Engl J Med. 2012;367:192–200.

4. Lubs HA, Stevenson RE, Schwartz CE. Fragile X and X-linked intellectual disability: forty decades of discovery. Am J Hum Genet. 2012;90:579–90.

5. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Mohseni M, Puttmann L, Vahid LN, Jensen C, Mohab LA, Hennies BC, Fultesi M, Ulich T, Weissmann R, Duan HH, Wang J, Wrogemann K, Hadiavi V, Lipkowitz B, Esmaeili-Niak S, Wierzbicki D, Kariminejad R, Frinxiaud-Sag D, Cohen M, Fastati J, Bost L, Moghavvghi F, Herzig B, Dehghani A, Rajab A, Banaian MJ, Hoffer J, Falah M, Musante L, Kalscheur V, Ullmann R, Kuss AW, Tschach H, Ahraki K, Ropers HH. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2011;478:57–63.

6. Ropers HH. Genetics of early onset cognitive impairment. Annu Rev Genomics Hum Genet. 2010;11:161–87.

7. Visseres LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wiekamp K, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. A de novo paradigm for mental retardation. Nat Genet. 2010;42:1109–12.

8. Rauch A, Wierzbicki D, Graf E, Wieland T, Schiavone L, Krusemeyer E, Cazzaniga G, Knappe R, Hatt H, Grigull S, Leunig F, Hoischen A, Kramer V, Deben H, Kostuzzi C, Hofmann S, Tschach H, Ahraki K, Ropers HH. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2011;478:57–63.

9. Gilissen CH-KJ, Thung DT, van de Vorst M, Willems MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Kirk D, Stratford S. The epidemiology of mental retardation. J Med Genet. 2008;45:176–83.

10. Lubs HA, Stevenson RE, Schwartz CE. Fragile X and X-linked intellectual disability: forty decades of discovery. Am J Hum Genet. 2012;90:579–90.

11. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Mohseni M, Puttmann L, Vahid LN, Jensen C, Mohab LA, Hennies BC, Fultesi M, Ulich T, Weissmann R, Duan HH, Wang J, Wrogemann K, Hadiavi V, Lipkowitz B, Esmaeili-Niak S, Wierzbicki D, Kariminejad R, Frinxiaud-Sag D, Cohen M, Fastati J, Bost L, Moghavvghi F, Herzig B, Dehghani A, Rajab A, Banaian MJ, Hoffer J, Falah M, Musante L, Kalscheur V, Ullmann R, Kuss AW, Tschach H, Ahraki K, Ropers HH. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2011;478:57–63.

12. Ropers HH. Genetics of early onset cognitive impairment. Annu Rev Genomics Hum Genet. 2010;11:161–87.

13. Visseres LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wiekamp K, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. A de novo paradigm for mental retardation. Nat Genet. 2010;42:1109–12.

14. Rauch A, Wierzbicki D, Graf E, Wieland T, Schiavone L, Krusemeyer E, Cazzaniga G, Knappe R, Hatt H, Grigull S, Leunig F, Hoischen A, Kramer V, Deben H, Kostuzzi C, Hofmann S, Tschach H, Ahraki K, Ropers HH. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2011;478:57–63.

15. Gilissen CH-KJ, Thung DT, van de Vorst M, Willems MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Kirk D, Stratford S. The epidemiology of mental retardation. J Med Genet. 2008;45:176–83.
Eng CM. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013;369:1502–11.

Amelri K, Manzoori M, Boucher P, Hogenesch JB, Pinson R, Shendure J, Eichler EE, Vissers LE, Menten B, Collard MW, de Vries BB. Mutations Affecting the ELP4, SRPX2, and interacting genes in typical and atypical rolandic epilepsy. *Am J Hum Genet* 2013;93:464–75.

Koide H, Ishida Y, Shinkawa K, Saito K, Minagawa T, Yamasaki K, Sato H, Tanaka T. Two distinct neurofibromatosis type 1 patients with TRPV1 mutation in exons 3 and 4. *Am J Med Genet B Neuropsychiatr Genet* 2013;164A:1565–70.

Kohno S, Okajima S, Kato T, Yagyu Y, Kageyama R, Takahashi M, Mariko T, Shinagawa Y, Nishihara H, Kojima S, et al. Novel missense variant in the DMD gene causes muscular dystrophy. *Am J Med Genet B Neuropsychiatr Genet* 2013;164A:1560–5.

Brophy MJ, Terech J, Leach S, Paoletti R, Williams J, et al. The role of the coatomer protein complex type I gamma subunit in the cellular localization and function of the coatomer protein complex type I alpha subunit. *Hum Mol Genet* 2016;25:1776–86.

Holt R, Webster M, Flanagan J, et al. Identification of a novel missense mutation in the MECP2 gene associated with Rett syndrome in a Chinese population. *Hum Genet* 2012;131:1025–9.

Eichler EE, Vissers LE, Menten B, Collard MW, de Vries BB. Mutations Affecting the ELP4, SRPX2, and interacting genes in typical and atypical rolandic epilepsy. *Am J Hum Genet* 2013;93:464–75.

Yvon Y, Robert C, Fauconnier A, et al. Neurocognitive, behavioral and somatic phenotypes associated with new missense, nonsense and small deletion mutations in the ANO1 gene. *Hum Genet* 2016;135:1639–50.

Akagi T, Wakiyama M, Nakamura T, et al. Whole-exome sequencing identifies a novel missense mutation in the DNM2 gene associated with hereditary nonfamilial adult-onset hypoglycemia. *Hum Genet* 2016;135:1609–18.

Carrilho I, Chabrol B, Champion MP, Coldwell J, Clayton P, Donner E, Gervais A, Haeusler JK, Huebner K, et al. Frequency and phenotypic spectrum of atypical and typical X-linked mental retardation. *Eur J Hum Genet* 2016;24:933–41.

Molloy KM, D’Agostino JM, Feng Z, Guo X, Kujal J, Uusimaa E, Shungu J, Lin B, et al. Whole-genome association study of very low birthweight reveals novel and gene interaction signals. *Sci Rep* 2016;6:30422.

Connor G, Prentice J, Shriver MD, et al. Whole-genome association study of type 1 diabetes highlights genetic overlap with type 2 diabetes and obesity. *Nat Genet* 2016;48:1413–20.

Zhang D, Liu J, Yang Q, et al. The role of RANBP3 in erythroid maturation and erythroleukemia. *Blood* 2016;127:1592–8.

Huang WH, Hsu WJ, Lee NC, et al. Novel missense andframeshiftmutations associated with intestinal diarrhea and aplastic anemia. *Hum Genet* 2016;135:1707–13.

Klinge S, Biliot J, Amsellem D, et al. Long-range organization of the human genome: Implications for the clinical interpretation of genomic sequencing studies. *Hum Genet* 2016;135:1963–76.

Kaspar AB, Finkielman RD, et al. Whole-exome sequencing identifies a novel missense mutation in the CYBB gene associated with chronic granulomatous disease. *Hum Genet* 2016;135:1837–46.

Mendez-Negrete MA, et al. Whole-exome sequencing identifies a novel missense mutation in the MAF gene associated with fulminant hepatic failure. *Hum Genet* 2016;135:1813–21.

Srivastava S, Chen SC, et al. Whole-exome sequencing identifies a novel missense mutation in the Olfm4 gene associated with primary ciliary dyskinesia. *Hum Genet* 2016;135:1831–4.

Dhawan V, Ramesh V, et al. Whole-exome sequencing identifies a novel missense mutation in the KIF20A gene associated with primary ciliary dyskinesia. *Hum Genet* 2016;135:1823–31.

Liu H, Xu W, Wang J, et al. Whole-exome sequencing identifies a novel missense mutation in the ARID1B gene associated with anaplastic large cell lymphoma. *Hum Genet* 2016;135:1789–97.

Kaspar AB, Finkielman RD, et al. Whole-exome sequencing identifies a novel missense mutation in the CYBB gene associated with chronic granulomatous disease. *Hum Genet* 2016;135:1837–46.

Klinge S, Biliot J, Amsellem D, et al. Long-range organization of the human genome: Implications for the clinical interpretation of genomic sequencing studies. *Hum Genet* 2016;135:1963–76.

Klinge S, Biliot J, Amsellem D, et al. Long-range organization of the human genome: Implications for the clinical interpretation of genomic sequencing studies. *Hum Genet* 2016;135:1963–76.

Klinge S, Biliot J, Amsellem D, et al. Long-range organization of the human genome: Implications for the clinical interpretation of genomic sequencing studies. *Hum Genet* 2016;135:1963–76.
Salbert B, Lacassie Y, Biser A, McDonald-McGinn DM, Zackai EH, Deardorff MA, Shaikh TH, Haan E, Friend KL, Fichera M, Romano C, Gez J, DeLisi LE, Sebat J, King MC, Shaffer LG, Eichler EE. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* 2010;42:203–9.

Leblond CS, Heinrich J, Delorme R, Proepper C, Betancur C, Huguet G, Konyukh M, Chaste P, Ey E, Rastarn M, Anckarsater H, Vanhala R, Curran S, Collier DA, Bolton P, Chiocchetti A, Klauck SM, Pouktka F, Freitag CM, Vanhala R, Kopp M, Ducetis E, Bacchelli E, Minopoli F, Ruta I, Battaglia A, Maestrini E, Sequiera AF, Oliveira B, Vicente A, Oliveira G, Finto D, Scherer SW, Zelenika D, Delapeine M, Lathrop M, Carrieri D, Guinchat V, Devillard F, Assouline B, Mouren MC, Leboyer M, Gillberg C, Boeckers TM, Bourgeron T. Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet* 2012;8:e1002521.

Schaaf CP, Sabo A, Sakai Y, Crosby J, Muzny D, Hawes A, Akbar H, Varghese R, Boenvinkle E, Gibbs RA, Zoghbi HY. Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum Mol Genet* 2011;20:3366–75.

Mandel JL, Chelly J. Monogenic X-linked mental retardation: is it as frequent as currently estimated? The paradox of the ARX (Aristaless X) mutations. *Eur J Hum Genet* 2004;12:689–93.

Nava C, Lamari F, Heron D, Mignot C, Rastetter A, Keren B, Cohen D, Faucet A, Bouettier D, Gillieron M, Jacquette A, Whalen S, Alenjar A, Perisse D, Laurent C, Dupuits C, Gautier C, Gerard M, Huguet G, Calliet S, Leheup B, Leboyer M, Gillberg C, Delorme R, Bourgeron T, Brice A, Depienne C. Analysis of the chromosome X exome in patients with autism spectrum disorders identified novel candidate genes, including TMLHE. *Transl Psychiatry* 2012;2:e1179.

Tarpey PS, Smith R, Pleasance E, Whitley A, Edkins S, Hardy C, O’Meara S, Latimer C, Dicks E, Menzies A, Stephens P, Blow M, Greenman C, Xue Y, Tyler-Smith C, Thompson D, Gray K, Andrews J, Barthorpe S, Buck G, Cole J, Dunmore R, Jones D, Maddison M, Mironenko T, Turner R, Turell K, Varian J, West S, Widdas S, Wray P, Teague J, Butler A, Jenkinson A, Jia M, Richardson D, Shepherd H, Wooster R, Tejada MI, Martinez F, Carvill G, Goliath R, de Brouwer AP, van Bokhoven H, Van Esch H, Chelly J, Raynaud M, Rogers HH, Abdi FE, Srivastava AK, Cox J, Luu Y, Malila U, Moon J, Parna J, Mohammed S, Tolmie JL, Sourbridge C, Corbett M, Gardner A, Haan E, Rujirabanjerd S, Shaw M, Vandelevre L, Fullston T, Easton DF, Boyle J, Partington M, Hackett A, Field M, Skinner C, Stevenson RE, Bobrow M, Turner G, Schwartz CE, Gez J, Raymond FL, Futeal FA, Stratton MR. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 2009;41:535–43.

Simpson DA, Clark GR, Alexander S, Silvestri G, Willoughby CE. Molecular diagnosis for heterogeneous genetic diseases with targeted high-throughput DNA sequencing applied to retinitis pigmentosa. *J Med Genet* 2011;48:145–51.

Brownstein Z, Friedman LM, Shahin H, Oror-Karni V, Kol N, Abu Rayyan A, Parafetall T, Lev D, Shalev S, Frydman M, Davidov B, Shohat M, Rahile M, Lieberman S, Levy-Lahad E, Lee MK, Shomron N, King MC, Walsh T, Kanaan M, Avraham KB. Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. *Genome Biol* 2011;12:R88.

Huang L, Jolly LA, Willis-Owen S, Gardner A, Kumar R, Douglas E, Sourbridge C, Wieczorek D, Tschach A, Cohen M, Hackett A, Field M, Froyen G, Hu H, Haas SA, Rogers HH, Kalscheuer VM, Corbett MA, Gez J. A noncoding, regulatory mutation implicates HCF1 in nonsyndromic intellectual disability. *Am J Hum Genet* 2012;91:694–702.

Ribbons RJ, Wada T, Fisher CA, Malik N, Mitson MJ, Steensma DP, Fryer A, Goudie DR, Krantz ID, Traeger-Synodinos J. Mutations in the chromatin-associated protein ATRX. *Hum Mutat* 2008;29:796–802.