Review

Genetics of recessive cognitive disorders

Luciana Musante and H. Hilger Ropers

Max Planck Institute of Molecular Genetics, Berlin, Germany

Most severe forms of intellectual disability (ID) have specific genetic causes. Numerous X chromosome gene defects and disease-causing copy-number variants have been linked to ID and related disorders, and recent studies have revealed that sporadic cases are often due to dominant *de novo* mutations with low recurrence risk. For autosomal recessive ID (ARID) the recurrence risk is high and, in populations with frequent parental consanguinity, ARID is the most common form of ID. Even so, its elucidation has lagged behind. Here we review recent progress in this field, show that ARID is not rare even in outbred Western populations, and discuss the prospects for improving its diagnosis and prevention.

ID: a major unsolved problem of healthcare

Early-onset cognitive impairment, commonly referred to as mental retardation or, more recently, ID [1], is defined as a disability 'characterized by significant limitations both in intellectual functioning and in adaptive behavior', and which 'originates before the age of 18' [2] with an IQ below 70 (= IQ 100 – 2SD) which is generally considered to be the threshold for ID. According to this definition, ID is estimated to affect 1–3% of Western populations [3] but is significantly more common elsewhere, with malnutrition, cultural deprivation, poor healthcare, and parental consanguinity as predisposing factors. Worldwide, ID is a major socioeconomic problem, the most costly of all diagnoses listed in the International Classification of Diseases (ICD10, http://www.cdc.gov/nchs/icd/icd10.htm), and the most frequent reason for referral to genetic services [4]. ID may be the only clinical symptom or it may be part of a clinically recognizable syndrome, but specific clinical features will often only be apparent when comparing several patients [5], and a sharp distinction between syndromic and non-syndromic forms (NS-ID) is not possible.

Most autosomal recessive gene defects are still unknown

Since 1991, the year when common fragile X syndrome was elucidated, more than 100 X-linked gene defects have been implicated in ID, as reported and reviewed elsewhere [6,7]. During the past decade numerous *de novo* and recurrent copy-number variants (CNVs) have been identified that cause or predispose to ID [8] and more recently, sequencing of affected individuals and their healthy parents has indicated that in sporadic patients, *de novo* basepair changes are another important cause of ID ([9–11] and references therein). By contrast, research into autosomal recessive ID (ARID) has lagged behind, possibly because in Western societies where most of the genetic research takes place, families are usually small, which has hampered mapping and identification of the underlying gene defects. This problem has been partly overcome by the introduction of high-throughput DNA sequencing techniques (Box 1). However, it has been shown that ARID is extremely heterogeneous, that the total number of ARID genes may run into the thousands (reviewed in [4]), and that the vast majority of these are still unknown.

Homozygosity mapping in consanguineous families

Homozygosity (or autozygosity) mapping in consanguineous families is the strategy of choice for mapping genes for recessive disorders in the human genome [12] (Box 1). Before 2002, virtually nothing was known about the molecular causes of ARID and, until 2006, no more than three genes for non-syndromic ARID had been identified, all by microsatellite-based homozygosity mapping in large consanguineous families and subsequent mutation screening of functionally plausible positional candidate genes [4] (Table 1).

The first large study employing single-nucleotide polymorphism (SNP) arrays to map ID genes [13] identified single homozygous linkage intervals in 8 of 76 consanguineous Iranian families with two or more affected children. None of these intervals overlapped, indicating that ARID is highly heterogeneous. This was confirmed by subsequent studies [14,15]. Thus, in contrast to non-syndromic recessive deafness, where 50% of the patients have mutations in a single gene (reviewed in [16]), these studies did not identify any frequent forms of ARID.

Many of the homozygous intervals in these families were large, which meant that hundreds of genes often had to be screened to identify the causative mutation. Nevertheless, systematic Sanger sequencing has led to the identification of numerous novel genes for non-syndromic ARID (Table 1) ([4] and references therein; [17,18]).

Corresponding author: Ropers, H.H. (ropers@molgen.mpg.de).
Keywords: autosomal recessive ID; homozygosity mapping; next-generation sequencing; healthcare.

0168-9525/$ – see front matter
 http://dx.doi.org/10.1016/j.tig.2013.09.008

Box 1. Traditional strategies to map ARID genes

Screening for disease-associated CNVs by array-comparative genomic hybridization (a-CGH)

CNVs are structural variations in the genome which consist in gains and losses of large chunks of DNA sequence with a range in length from 1000 bp to 5 Mb (cytogenetic level of resolution). Because CNVs change the structure of the genome, their functional effect could crucially depend on whether they change the sequence or relative location of specific segments of genomic DNA.

Linkage mapping in multiple-affected families

Genetic linkage is the tendency whereby alleles at loci close to each other on a chromosome will be inherited together during meiosis because they are less likely to be separated by a crossover event. Conversely, if loci are far apart or on different chromosomes then recombination will occur by chance in 50% of meioses. The recombination fraction ranges from 0 (tight linkage) to 0.5 (no linkage) and is a measure of genetic distance. Linkage can be used to map disease genes by typing DNA markers (i.e., SNPs) and seeing if their alleles cosegregate with the disease phenotype.

Homozygosity mapping in consanguineous families

Consanguineous families are common in countries belonging to the 'consanguinity belt' that extends from Morocco to India, and in migrant communities now permanently resident in Western Europe, North America, and Australasia [96] (see also http://www.consang.net/). It is estimated that about 20% of the human population live in communities with a preference for consanguineous marriage and that at least 8.5% of children have consanguineous parents ([97] and references therein). Globally, the most common form of consanguineous union is between first cousins, who share 1/8 of their genes, and their progeny therefore show autozygosity at 1/16 of all loci. Conventionally, this is expressed as the coefficient of inbreeding (F), and for first-cousin offspring $F = 0.0625$ [98]. The children of consanguineous individuals will have more homozygous DNA than the offspring of an outbred marriage. This leads to an increased likelihood of rare, recessive disease-causing variants being inherited from a common ancestor via both maternal and paternal lineages. Homozygosity mapping is based on the fact that the affected offspring of consanguineous matings will not only be homozygous by descent for the causative gene defect, but also for flanking genetic markers located on the same chromosomal segment.

Box 2. Disease gene identification by NGS

In the past a traditional way to identify Mendelian disease genes was Sanger sequencing of candidate genes selected by positional mapping (i.e., linkage analysis, homozygosity mapping), by their relation to other genes responsible for similar phenotypes, or because the encoded proteins were known to be physiologically or functionally relevant to the disease in question. The introduction of NGS has revolutionized the genetic dissection of monogenic diseases, allowing the identification of gene defects underlying ID in familial cases even where linkage analysis would be impossible due to insufficient family information (size of the family, number of affected per family, etc.) as well as in sporadic cases and encompassing diverse models of inheritance. Moreover, it can be applied to the detection of CNVs. Few major NGS platforms exist (reviewed in [108]). Although they use different enzymology, chemistry, high resolution optics, hardware and software, nevertheless they share some commonalities – they generally start with fragmented genomic DNA, ligated with platform specific linker, then selectively amplified by PCR, ready for massively parallel sequencing resulting in millions of short reads. NGS can be applied to sequencing of the entire human genome (referred to as whole-genome sequencing, WGS), to the entire protein-coding sequences (known as whole-exome sequencing, WES), and also to a subset of genomic regions (i.e., exons within the homozygous loci or linkage intervals) or to a subset of target genes. Despite the advantage of NGS technologies compared to previous methods, including increased speed and reduced costs, the major challenge resides now in the interpretation of the large number of variants identified. It will be crucial to develop strategies for disease variant prioritization, including robust bioinformatics procedures to filter the relevant changes. This process could take advantage also of the development of databases of genetic variants present in affected and healthy individuals.

Next-generation sequencing (NGS): a new dimension in the elucidation of ARID

The introduction of high-throughput NGS techniques (Box 2) has revolutionized the genetic dissection of ID and the identification of gene defects underlying ARID. *TECR* was the first gene for which a causative homozygous variant was identified by whole-exome enrichment and

Table 1. NS-ARID genes identified before the NGS era (2002–2011)

Gene[a]	HGNC ID	Number of families reported	Ethnicity	Mutation[b]	Disorder[c]	OMIM[d]	First description
PRSS12	9477	2	Algerian	delACGT 1350-1353	MRT1, NS-ID	#249500	[99]
CRBN	30185	1	Closed population (North America)	R419X	MRT2, NS-ID	#607417	[100]
CC2D1A	30237	9	Israeli Arab	G408fsX437	MRT3, NS-ID	#608443	[101]
GRIK2	4580	1	Iranian	del/inv, Ex7–11	MRT6, NS-ID	#611092	[102]
TUSC3	30242	5	Iranian French Pakistani Italian	del120 Kb; Q55X N263TfsX300 del170 kb del203 kb	MRT7, NS-ID	#611093	[103,104]
TRAPPC9	30832	7	Israeli Arab Tunisian Pakistani Iranian Syrian Italian	R475X R570X R475X; c.1024+1G>T L772WfsX7 R475X T951YfsX17	MRT13, NS-ID	#613192	[105–107]
ZC3H14	20509	1	Iranian	R154X	NS-ID		[17]
MED23	2372	1	Algerian	R617Q	NS-ID	#614249	[18]

[a]Gene symbol approved by Human Gene Nomenclature Committee, HGNC (http://www.genenames.org/).

[b]Abbreviations: c, coding region; del, deletion; fs, frameshift; inv, inversion; X, stop codon.

[c]MRT, mental retardation, autosomal recessive, phenotypic series, OMIM, Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim); NS-ID, non-syndromic intellectual disability.

[d]OMIM number (#), phenotypic description, molecular basis known, version 9 October 2013.

Table 2. NS-ARID (candidate) genes identified by NGS (since 2011)[a]

Gene	HGNC ID	Number of families reported	Ethnicity	Mutation	Disorder	OMIM	First description
ADK	257	1	Iranian	H324R	NS-ID ASD		[22]
ADRA2B	282	1	Iranian	R440G	NS-ID		[22]
ASCC3	18697	1	Iranian	S1564P	NS-ID		[22]
ASCL1	738	1	Iranian	A41S	NS-ID		[22]
C11orf46	26798	1	Iranian	R236H	NS-ID		[22]
TTI2	26262	1	Iranian	P367L	NS-ID		[22]
RABL6	24703	1	Pakistani	A562P	NS-ID		[22]
CASP2	1503	1	Iranian	Q392X	NS-ID		[22]
CCNA2	1578	1	Iranian	Splice site	NS-ID		[22]
COQ5	28722	1	Iranian	G118S	NS-ID		[22]
CRADD	2340	5	Old Order Amish and Mennonite	G128R	MRT34, NS-ID	#614499	[43]
EEF1B2	3208	1	Iranian	Splice site	NS-ID		[22]
ELP2	18248	2	Lebanese Iranian	T555P R462L	NS-ID		[22]
ENTPD1	3363	1	Iranian	Y65C	NS-ID		[22]
FASN	3594	1	Iranian	R1819W	NS-ID		[22]
HIST3H3	4778	1	Omani	R130C	NS-ID		[22]
INPP4A	6074	1	Iranian	D915fsX	NS-ID		[22]
KIAA1033	29174	1	Omani	P1019R	NS-ID		[109]
MAN1B1	6823	3 1 1	Pakistani Iranian	E397K W473X R334C	MRT15, NS-ID	#614202	[22,41]
NDST1	7680	1	Iranian	R709Q	NS-ID		[22]
PECR	18281	1	Iranian	L57V	NS-ID ASD		[22]
PRMT10	25099	1	Iranian	G189R	NS-ID		[22]
PRRT2	30500	1	Iranian	A214fsX	NS-ID		[22]
RALGDS	9842	1	Iranian	A706V	NS-ID		[22]
RGS7	10003	1	Iranian	N304fsX	NS-ID ASD		[22]
SCAPER	13081	1	Iranian	Y118fsX	NS-ID		[22]
ST3GAL3	10866	2	Iranian	A13D D370Y	MRT12, NS-ID	#611090	[39]
TECR	4551	15	Endogamic population; Hutterites	P182L	MRT14, NS-ID	#614020	[19]
TRMT1	25980	1	Iranian	I230fsX	NS-ID		[22]
UBR7	20344	1	Iranian	N124S	NS-ID ASD		[22]
ZCCHC8	25265	1	Iranian	L90X	NS-ID		[22]
ZNF526	29415	2	Iranian	R459Q Q539H	NS-ID		[22]

[a]Abbreviation: ASD, autism spectrum disorder; other abbreviations are given in Table 1 legend.

sequencing (WES) of a large consanguineous family with NS-ID [19], and a missense mutation in this gene was recently found to be a common cause of NS-ID in Hutterites [20]. *TECR* codes for *trans*-2,3-enoyl-CoA reductase (also referred to as synaptic glycoprotein 2), which reduces *trans*-2,3-stearoyl-CoA to stearoyl-CoA of long and very long chain fatty acids (VLCFA). Perturbations of VLCFA metabolism have also been observed in other neurological disorders such as adrenoleukodystrophy and Zellweger syndrome, and mutations affecting FACL4, which is involved in the degradation of VLCFA and the production of key intermediates in the synthesis of complex lipids, are known to cause X-linked ID [21].

More recently, a large study highlighted the extraordinary potential of NGS for unraveling the molecular basis of ARID [22]. Instead of performing WES, these authors opted for the enrichment and sequencing of exons from homozygous linkage intervals in consanguineous Iranian families. In 78 of 136 families investigated a single, apparently disease-causing sequence variant was identified. Of these families, 26 had homozygous mutations in 23 genes previously implicated in ID or related neurological disorders and, in addition, single homozygous mutations were found in 50 novel candidate genes for ARID, mostly in patients with apparently 'pure' or NS-ID (Table 2). Follow-up studies have revealed additional clinical symptoms in patients with mutations involving the same genes, thereby confirming their postulated role in ID, but also illustrating the clinical variability of these gene defects.

It is noteworthy that, in about 40% of the families studied, potentially causative gene defects could not be identified. In populations where parental consanguinity is common, not all recessive conditions are due to autozygous changes, which were the only target of this study; other

defects including compound heterozygosity or mutations in intronic, promoter, or other non-coding sequences could not be detected by this approach. In other families, pathogenic changes may have been overlooked due to overly-stringent filtering of sequence variants, including all synonymous changes.

In outbred populations, most patients with recessive forms of ID or related disorders will be sporadic cases. ID families are mostly non-consanguineous and only a small proportion have multiple affected siblings [23]. Recently, the first systematic WES study including 19 such families revealed compound heterozygous frameshift changes in the *DDHD2* gene, which encodes one of the three mammalian intracellular phospholipases A(1) [24], as well as pathogenic mutations in two known X-linked ID genes. Potentially pathogenic mutations, including three compound heterozygous and two homozygous changes, were identified in five candidate genes not previously implicated in ID [23]. Thus, the diagnostic yield of this study (42%) was only slightly inferior to one performed in consanguineous families (57%) [22], although it remains to be seen how many will be confirmed by validation studies.

Most novel candidates are bona fide ARID genes

Many of the recently reported novel candidate genes are very attractive candidates because of their synapse- or brain-specific function; others involve basic cellular processes which have been repeatedly implicated in ID, such as DNA transcription and translation, protein degradation, mRNA splicing, energy metabolism, or fatty-acid synthesis and turnover [22]. Conclusive proof for their indispensable role in the brain has been obtained for a growing number of these genes through the identification of additional mutations in unrelated families, studies in mouse or fly models, or by other means.

For example, mutations in the *LARP7* gene have now been observed in two unrelated families. *LARP7* encodes a negative transcriptional regulator of polymerase II genes, acting by means of the 7SK ribonucleoprotein (RNP) system [25]. After the first description [22], a second loss-of-function mutation in *LARP7* was described in a family from Saudi Arabia with primordial dwarfism, intellectual disability, and dysmorphic facial features [26].

In a consanguineous family with ID, facial dysmorphisms, and cataracts, a homozygous intragenic *CACNA1G* deletion was described that is predicted to remove at least 20 amino acids of CACNA1G, abolishing its function. CACNA1G is a T-type calcium channel with a crucial role in the generation of $GABA_B$ receptor-mediated spike and wave discharges in the thalamo-cortical pathway ([27] and references therein). A second homozygous *CACNA1G* mutation has been found that removes a single but apparently essential amino acid (F.S. Alkuraya, Riyadh, personal communication) in several members of a previously described Arab family with a severe syndromic form of ARID [28]. Moreover, a *de novo* deletion removing one copy of the *CACNA1G* gene has been found in a male patient [29] whose clinical features closely resembled that of Iranian patients with a homozygous intragenic *CACNA1G* deletion [22].

Mutations in the *NSUN2* gene have been identified in five unrelated consanguineous families. Together, these findings revealed the syndromic nature of this condition, which includes characteristic facial features and variable other clinical signs [30–32]. *NSUN2* encodes an RNA methyltransferase which methylates cytosine to 5-methylcytosine (m5C) at position 34 of intron-containing tRNA(Leu)(CAA) precursors [33]. A *Drosophila* model of this was generated by deleting the *NSUN2* ortholog, which resulted in severe short-term memory (STM) deficits, pointing to an important role of RNA methylation in cognition [30]. *NSUN2* is now the third RNA-methyltransferase gene linked to ID. Previously, *FTSJ1* (MRX9, MIM #309549) had been implicated in X-linked NS-ID [34], and recently *TRMT1*, which encodes a tRNA (G26) dimethyltransferase, was identified as a novel candidate gene for ARID [22].

ZC3H14, mutated in a consanguineous family with NS-ARID, is another gene whose indispensable role in the central nervous system has been supported by a *Drosophila* model. ZC3H14 is the human ortholog of the *Drosophila* Nab2 protein, which binds to polyadenylated mRNA and restricts the length of the poly(A) tail, and this protein was also found to be indispensable for normal behavior in the fly [17] (see Table 1). *ZC3H14* is a new member of the growing list of ID genes with a role in mRNA metabolism, including *FMRP, FMR2P, PQBP1, UFP3B, DYRK1A*, and *CDKL5* ([35] for review).

ARID is extremely heterogeneous and clinically variable

At the time of writing, 40 genes have been implicated in NS-ID (Tables 1 and 2). In 11 of these, apparently pathogenic mutations have been detected in more than one family. A mutation in the neurotrypsin gene (*PRSS12*) has been found in two apparently unrelated Algerian families with NS-ID (reviewed in [4] and references therein). A mutation in the *CC2D1A* gene, the product of which regulates expression of the serotonin receptor 1A gene in neuronal cells, had been identified in nine nuclear families and more recently in a Pakistani family ([36] and references therein). Other established ARID genes include *TUSC3* which is required for cellular Mg^{2+} uptake, trafficking protein particle complex 9 (*TRAPPC9*), and ST3 β-galactoside α-2,3-sialyltransferase 3 (*ST3GAL3*) ([4] and references therein; [37–40]), *MAN1B1* encoding an enzyme which functions in *N*-glycan biosynthesis [22,41], the transcriptional regulator *ZNF526*, and *ELP2* [22] which encodes a subunit of the RNA polymerase II elongator complex [42]. Finally, *CRADD* has been identified as new gene for NS-ARID in affected children from different Old Order Amish and Mennonite sibships [43]. *CRADD* codes for a caspase recruitment domain and death domain-containing adaptor protein that activates caspase 2, a novel candidate gene for NS-ARID [22], and is required for neuronal apoptosis [44].

For many of the recently described gene defects that give rise to ARID the clinical picture has turned out to be complex and variable. *ADK* deficiency may lead to NS-ARID or present with severe developmental delay, persistent hypermethioninemia, and mild liver dysfunction [45], and *KIF7* mutations have been reported in two different, clinically distinguishable ID-malformation syndromes ([46] and references therein). ARID genes have also been implicated in conditions that are apparently unrelated to

ID, pointing to pleiotropic functions of these genes. For example, overexpression of the fatty acid synthase *FASN*, a strong positional and functional candidate gene for ARID [22], predisposes to leiomyomatosis [47], and homozygous inactivation of *FTO*, which encodes an RNA demethylase and has been previously implicated in obesity [48], has been shown to result in severe developmental delay with malformations [49].

A role for recessive factors in epilepsy, autism, and other psychiatric disorders?

ID is frequently associated with psychiatric and/or neurological disorders (reviewed in [50]). Based on the International Classification of Diseases (ICD, 10th revision) it has been estimated that between 14% and 39% of individuals with ID present with comorbid psychiatric diagnoses.

Epilepsy is among the most frequently associated disorders [50], with a frequency ranging from 5.5% to 35%, which is similar to the 20–27% reported by population-based studies of children with epilepsy and some degree of ID ([51] and references therein). In patients with mild to moderate ID its frequency is 15%, but it may exceed 30% if the ID is severe or profound [52]. Moreover, epilepsy is seen in about half of the X-linked ID syndromes (reviewed in [53]). A number of well-known genetic disorders share ID, epilepsy, and autism as prominent clinical features, including tuberous sclerosis, Rett syndrome, and fragile X [54].

In recent years, the contribution of structural genome variation to epilepsy has become increasingly evident. The most common CNVs associated with epilepsy, at 15q13.3, 15q11.2, and 16p13.11, also confer susceptibility for learning disabilities (reviewed in [8] and references therein) suggesting that common genetic factors could have a causative role. NGS has also been instrumental in identifying genes for recessive syndromes encompassing epilepsy and ID. One of the earliest applications of this technology was the identification of homozygous and compound heterozygous changes in the *TBC1D24* gene, which encodes a Rab GTPase activator, in an Arab family with seizures and ID [55], and in an Italian family with infantile myoclonic epilepsy (MIM #605021) [56], respectively. Since then, two additional families with recessive mutations in *TBC1D24* have been described with early infantile epileptic encephalopathy 16 (MIM #615338) [57,58].

Recently, homozygous frameshift mutations in the *PRRT2* (proline-rich transmembrane protein 2) gene have been identified in two families with ID and epilepsy in infancy [22,59], whereas heterozygous truncating and missense mutations were shown to cause dominant infantile epilepsy (MIM #605751) and episodic dyskinesia (MIM #128200) ([60] and references therein). These findings again highlight the stunning clinical variability of mutations involving the same gene.

In patients with ID autistic signs are also common, and most patients with autism have some degree of cognitive impairment [61]. Numerous genetic defects have been implicated in ID and autism, including mutations in X-linked genes (e.g., *NLG3, NLG4* [62,63], *TMLHE* [64], and CNVs (reviewed in [8]) or apparently dominant *de novo* mutations ([65] and references therein), and it is likely that the strong comorbidity between ID and other major psychiatric disorders [66] is also due, at least in part, to shared genetic factors (e.g., [67–69]).

In several genes for autosomal recessive ID, including *KDM6B* (lysine demethylase 6B), *MED13L*, which encodes for a subunit of the mediator complex, and nudE nuclear distribution E homolog 1 (*NDE1*) [22,70], dominant *de novo* mutations or loss of one entire gene copy have been described in autistic patients [71,72]. Nevertheless, there is little direct evidence for a causative role of recessive gene defects in autism and other psychiatric disorders.

In part this may be due to the focus of autism and schizophrenia research on common genetic risk factors (e.g., [73]) and, recently, on dominant *de novo* mutations (e.g., [71,74–76]) which may account for about 20% the sporadic cases [65]. Array CGH studies in consanguineous families have identified several homozygous deletions encompassing autism candidate genes, one of which was also found to be mutated in a non-consanguineous family with autism spectrum disorder (ASD) [77]. In simplex ASD families, affected individuals with IQ<70 have longer homozygous segments in their genome than unaffected siblings, but probands with an IQ>70 do not show this excess. Thus, long stretches of homozygosity may confer susceptibility to autism with low IQ or to low IQ alone [78].

Homozygous, compound heterozygous, or homozygous hypomorphic mutations in disease genes are known to associate with monogenic autosomal or X-linked recessive neurodevelopmental disorders, and potentially causative mutations in candidate genes were found in consanguineous and outbred ASD families by WES [79,80]. This finding, together with the identification of a twofold increase in rare complete knockout mutations in ASD patients compared to controls [81], provide convincing evidence that autosomal recessive gene defects play a role in autism, but their frequency is still unknown.

Finally, consanguinity has also been suggested as a risk factor for bipolar disorder and schizophrenia [82,83], but even less is known about the contribution of autosomal recessive mutations to the pathogenesis of these diseases.

How frequent are recessive forms of ID?

In the small families of outbred Western societies, most patients with recessive forms of ID or related disorders will be sporadic cases. If couples with offspring have two children on average, which is close to the actual situation in Europe ([84]; M. Kreyenfeld, Rostock, personal communication), only one of four patients will have an affected sibling and will be identified as a familial case. In Central Europe, between 3.3 and 6% of patients with ID referred to genetic services are familial cases [10,23]. Taken at face value, this suggests that recessive forms of ID account for 13–24% of the cases in Europe. However, this may be an overestimate because it is based on the assumption that parents with a single affected child will be equally likely to seek genetic advice as parents with two or more affected children, which is probably not true.

Given the low rate of parental consanguinity in developed countries, most patients with ARID are expected to be compound heterozygotes carrying two different disease-causing alleles [85]. This is in keeping with a recent study

focusing on dominant *de novo* mutations in sporadic ID [86]. No homozygous disease-causing mutation was found in 51 sporadic patients, but some carried two allelic and probably pathogenic mutations in functional candidate genes, suggesting that a minor proportion of the cases may be due to ARID. However, the true proportion of ARID must be higher because familial cases and consanguineous families were not included in this study, some compound heterozygotes may have been overlooked because they are more difficult to detect by NGS, and mutations in non-coding DNA have not been taken into consideration. Detectable and submicroscopic chromosomal rearrangements account for approximately 25% of all individuals with severe ID, and X-linked factors are thought to be responsible for 10–12%. *De novo* mutations have been found in 16 and 31%, respectively, of sporadic patients [11,86], but their true frequency may be even higher. Taken together, in outbred populations ARID may account for about 10–20% of the cases, which leaves room for oligogenic/polygenic forms of ID, which have been the subject of a recent review [87].

In populations where parental consanguinity is common, autosomal recessive gene defects must be an even more important cause of ID. In families from the Middle East, autosomal recessive disorders were found to be almost threefold more frequent among inbred as among non-inbred cases [88]. In Jordan, autosomal recessive inheritance was observed in 32% of the families counseled and, of the ~27% sporadic cases without a definite diagnosis, 30% were also ascribed to autosomal recessive gene defects [89]. Thus, in these countries, ARID should be the most common genetic cause of ID – and a particularly promising target for diagnosis and prevention.

Implications for research and healthcare

Despite the remarkable progress in the elucidation of autosomal recessive forms of ID, it is likely that the several hundred genes already implicated in syndromic or non-syndromic ARID (see [90] and references therein) and related disorders are vastly outnumbered by the many ARID genes still waiting to be found. Considering that on the X chromosome alone, which carries 4% of all human genes, already more than 100 ID genes have been identified [6,7], there should be at least 2500 autosomal ID genes, and most of the novel forms of ID should be autosomal recessive, which is supported by functional considerations and evidence from model organisms.

NGS in families with two or more affected individuals has proven to be an extraordinarily effective approach for identifying novel recessive causes of ID, and international collaborations including the GENCODYS consortium (http://www.gencodys.eu/) have set out to identify the molecular causes of ARID in a systematic fashion. Although autozygosity mapping followed by targeted exon sequencing [22] is a successful and cost-effective strategy for finding causative gene defects in consanguineous families, it will only detect homozygous mutations. However, in families from Western industrialized countries, compound heterozygous mutations are common, and even in countries with frequent parental consanguinity, compound heterozygosity is not rare (H. Najmabadi, Tehran, personal communication). This argues for using WES as a more comprehensive strategy to elucidate novel causes of ARID, even though, as with targeted exon sequencing, it will miss most non-exonic mutations. Intronic changes [91] and mutations in non-coding regulatory sequences are only detectable by whole-genome sequencing (WGS), and another advantage of WGS is its more even coverage. This is why WGS does not require very high sequencing depths, and it may soon become an affordable alternative to WES.

Increasingly, WES has been proposed as a comprehensive diagnostic tool for detecting mutations in patients with ID and related disorders [92,93]. In a diagnostic setting, targeted NGS-based tests encompassing all genes implicated in ID or related disorders could be equally useful, but much cheaper and easier to read, and they will not yield any unsolicited results. Targeted tests of this kind have been developed for a variety of genetically heterogeneous conditions such as deafness and blindness, and a broad test for severe recessive childhood diseases is already routinely employed in healthcare [94,95].

Concluding remarks

After having been disregarded for a long time, recessive gene defects are being discovered at a rapid pace as important causes of ID. Comprehensive and affordable tests to rule out all known forms of ARID will have a major effect on the diagnosis and prevention of ID, not only in developing countries where parental consanguinity is common but also elsewhere.

Acknowledgments

We thank Hossein Najmabadi and Kimia Kahrizi, Hao Hu, Masoud Garshasbi, Andreas Kuss, Wei Chen, and Thomas Wienker for their essential contributions to our past and ongoing ARID research, and Gabriele Eder for help with the preparation of the manuscript. This work was supported by the Max Planck Society and by the European Commission Framework Program 7 (FP7) project GENCODYS, grant no 241995 (coordinator: Hans van Bokhoven, Nijmegen).

References

1 Schalock, R.L. *et al.* (2007) The renaming of mental retardation: understanding the change to the term intellectual disability. *Intellect. Dev. Disabil.* 45, 116–124
2 American Psychiatric Association (2000) *Diagnostic and Statistic Manual of Mental Disorders* (4th Revision: DSM-IV-TR), American Psychiatric Association
3 Leonard, H. and Wen, X. (2002) The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment. Retard. Dev. Disabil. Res. Rev.* 8, 117–134
4 Ropers, H.H. (2010) Genetics of early onset cognitive impairment. *Annu. Rev. Genomics Hum. Genet.* 11, 161–187
5 Kleefstra, T. and Hamel, B.C. (2005) X-linked mental retardation: further lumping, splitting and emerging phenotypes. *Clin. Genet.* 67, 451–467
6 Lubs, H.A. *et al.* (2012) Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* 90, 579–590
7 Piton, A. *et al.* (2013) XLID-causing mutations and associated genes challenged in light of data form large-scale human exome sequencing. *Am. J. Hum. Genet.* 93, 368–383
8 Mefford, H.C. *et al.* (2012) Genomics, intellectual disability, and autism. *N. Engl. J. Med.* 366, 733–743
9 Veltman, J.A. and Brunner, H.G. (2012) De novo mutations in human genetic disease. *Nat. Rev. Genet.* 13, 565–575
10 Rauch, A. *et al.* (2006) Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am. J. Med. Genet. A* 140, 2063–2074

11 de Ligt, J. *et al.* (2012) Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* 367, 1921–1929
12 Lander, E.S. and Botstein, D. (1987) Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236, 1567–1570
13 Najmabadi, H. *et al.* (2007) Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum. Genet.* 121, 43–48
14 Kuss, A.W. *et al.* (2011) Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum. Genet.* 129, 141–148
15 Abou Jamra, R. *et al.* (2011) Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. *Eur. J. Hum. Genet.* 19, 1161–1166
16 Duman, D. and Tekin, M. (2012) Autosomal recessive nonsyndromic deafness genes: a review. *Front. Biosci.* 17, 2213–2236
17 Pak, C. *et al.* (2011) Mutation of the conserved polyadenosine RNA binding protein, ZC3H14/dNab2, impairs neural function in *Drosophila* and humans. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12390–12395
18 Hashimoto, S. *et al.* (2011) MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science* 333, 1161–1163
19 Caliskan, M. *et al.* (2011) Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. *Hum. Mol. Genet.* 20, 1285–1289
20 Chong, J.X. *et al.* (2012) A population-based study of autosomal-recessive disease-causing mutations in a founder population. *Am. J. Hum. Genet.* 91, 608–620
21 Meloni, I. *et al.* (2002) FACL4, encoding fatty acid-CoA ligase 4, is mutated in nonspecific X-linked mental retardation. *Nat. Genet.* 30, 436–440
22 Najmabadi, H. *et al.* (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478, 57–63
23 Schuurs-Hoeijmakers, J.H.M. (2012) *Gene Identification in Intellectual Disability*, Radboud University Nijmegen
24 Schuurs-Hoeijmakers, J.H. *et al.* (2012) Mutations in DDHD2, encoding an intracellular phospholipase A(1), cause a recessive form of complex hereditary spastic paraplegia. *Am. J. Hum. Genet.* 91, 1073–1081
25 Markert, A. *et al.* (2008) The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* 9, 569–575
26 Alazami, A.M. *et al.* (2012) Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum. Mutat.* 33, 1429–1434
27 Zamponi, G.W. *et al.* (2010) Role of voltage-gated calcium channels in epilepsy. *Pflugers Arch.* 460, 395–403
28 Al-Owain, M. *et al.* (2011) An autosomal recessive syndrome of severe cognitive impairment, dysmorphic facies and skeletal abnormalities maps to the long arm of chromosome 17. *Clin. Genet.* 80, 489–492
29 Preiksaitiene, E. *et al.* (2012) A novel de novo 1.8 Mb microdeletion of 17q21.33 associated with intellectual disability and dysmorphic features. *Eur. J. Med. Genet.* 55, 656–659
30 Abbasi-Moheb, L. *et al.* (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 90, 847–855
31 Khan, M.A. *et al.* (2012) Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 90, 856–863
32 Martinez, F.J. *et al.* (2012) Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J. Med. Genet.* 49, 380–385
33 Brzezicha, B. *et al.* (2006) Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Res.* 34, 6034–6043
34 Freude, K. *et al.* (2004) Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. *Am. J. Hum. Genet.* 75, 305–309
35 Bardoni, B. *et al.* (2012) Intellectual disabilities, neuronal posttranscriptional RNA metabolism, and RNA-binding proteins: three actors for a complex scenario. *Prog. Brain Res.* 197, 29–51
36 Noor, A. *et al.* (2008) CC2D2A, encoding a coiled-coil and C2 domain protein, causes autosomal-recessive mental retardation with retinitis pigmentosa. *Am. J. Hum. Genet.* 82, 1011–1018
37 Khan, M.A. *et al.* (2011) A novel deletion mutation in the TUSC3 gene in a consanguineous Pakistani family with autosomal recessive nonsyndromic intellectual disability. *BMC Med. Genet.* 12, 56
38 Loddo, S. *et al.* (2013) Homozygous deletion in TUSC3 causing syndromic intellectual disability: a new patient. *Am. J. Med. Genet. A* 161, 2084–2087
39 Hu, H. *et al.* (2011) ST3GAL3 mutations impair the development of higher cognitive functions. *Am. J. Hum. Genet.* 89, 407–414
40 Marangi, G. *et al.* (2013) TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. *Eur. J. Hum. Genet.* 21, 229–232
41 Rafiq, M.A. *et al.* (2011) Mutations in the alpha 1,2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* 89, 176–182
42 Winkler, G.S. *et al.* (2001) RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes. *J. Biol. Chem.* 276, 32743–32749
43 Puffenberger, E.G. *et al.* (2012) Genetic mapping and exome sequencing identify variants associated with five novel diseases. *PLoS ONE* 7, e28936
44 Ribe, E.M. *et al.* (2012) Neuronal caspase 2 activity and function requires RAIDD, but not PIDD. *Biochem. J.* 444, 591–599
45 Bjursell, M.K. *et al.* (2011) Adenosine kinase deficiency disrupts the methionine cycle and causes hypermethioninemia, encephalopathy, and abnormal liver function. *Am. J. Hum. Genet.* 89, 507–515
46 Ali, B.R. *et al.* (2012) A mutation in KIF7 is responsible for the autosomal recessive syndrome of macrocephaly, multiple epiphyseal dysplasia and distinctive facial appearance. *Orphanet J. Rare Dis.* 7, 27
47 Eggert, S.L. *et al.* (2012) Genome-wide linkage and association analyses implicate FASN in predisposition to uterine leiomyomata. *Am. J. Hum. Genet.* 91, 621–628
48 Frayling, T.M. *et al.* (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316, 889–894
49 Boissel, S. *et al.* (2009) Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. *Am. J. Hum. Genet.* 85, 106–111
50 Oeseburg, B. *et al.* (2011) Prevalence of chronic health conditions in children with intellectual disability: a systematic literature review. *Intellect. Dev. Disabil.* 49, 59–85
51 Berg, A.T. and Plioplys, S. (2012) Epilepsy and autism: is there a special relationship? *Epilepsy Behav.* 23, 193–198
52 Prince, E. and Ring, H. (2011) Causes of learning disability and epilepsy: a review. *Curr. Opin. Neurol.* 24, 154–158
53 Stevenson, R.E. *et al.* (2012) Seizures and X-linked intellectual disability. *Eur. J. Med. Genet.* 55, 307–312
54 Leung, H.T. and Ring, H. (2013) Epilepsy in four genetically determined syndromes of intellectual disability. *J. Intellect. Disabil. Res.* 57, 3–20
55 Corbett, M.A. *et al.* (2010) A focal epilepsy and intellectual disability syndrome is due to a mutation in TBC1D24. *Am. J. Hum. Genet.* 87, 371–375
56 Falace, A. *et al.* (2010) TBC1D24, an ARF6-interacting protein, is mutated in familial infantile myoclonic epilepsy. *Am. J. Hum. Genet.* 87, 365–370
57 Guven, A. and Tolun, A. (2013) TBC1D24 truncating mutation resulting in severe neurodegeneration. *J. Med. Genet.* 50, 199–202
58 Milh, M. *et al.* (2013) Novel compound heterozygous mutations in TBC1D24 cause familial malignant migrating partial seizures of infancy. *Hum. Mutat.* 34, 869–872
59 Labate, A. *et al.* (2012) Homozygous c.649dupC mutation in PRRT2 worsens the BFIS/PKD phenotype with mental retardation, episodic ataxia, and absences. *Epilepsia* 53, e196–e199
60 Helbig, I. and Lowenstein, D.H. (2013) Genetics of the epilepsies: where are we and where are we going? *Curr. Opin. Neurol.* 26, 179–185

61 Fombonne, E. (2003) Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J. Autism Dev. Disord.* 33, 365–382
62 Jamain, S. *et al.* (2003) Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat. Genet.* 34, 27–29
63 Laumonnier, F. *et al.* (2004) X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am. J. Hum. Genet.* 74, 552–557
64 Celestino-Soper, P.B. *et al.* (2012) A common X-linked inborn error of carnitine biosynthesis may be a risk factor for nondysmorphic autism. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7974–7981
65 Jiang, Y.H. *et al.* (2013) Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am. J. Hum. Genet.* 93, 249–263
66 Morgan, V.A. *et al.* (2008) Intellectual disability co-occurring with schizophrenia and other psychiatric illness: population-based study. *Br. J. Psychiatry* 193, 364–372
67 Kirov, G. *et al.* (2009) Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum. Mol. Genet.* 18, 1497–1503
68 de Kovel, C.G. *et al.* (2010) Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 133, 23–32
69 Betancur, C. (2011) Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res.* 1380, 42–77
70 Guven, A. *et al.* (2012) Novel NDE1 homozygous mutation resulting in microhydranencephaly and not microlyssencephaly. *Neurogenetics* 13, 189–194
71 Iossifov, I. *et al.* (2012) De novo gene disruptions in children on the autistic spectrum. *Neuron* 74, 285–299
72 Ullmann, R. *et al.* (2007) Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum. Mutat.* 28, 674–682
73 Nicolson, R. and Szatmari, P. (2003) Genetic and neurodevelopmental influences in autistic disorder. *Can. J. Psychiatry* 48, 526–537
74 Sanders, S.J. *et al.* (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237–241
75 O'Roak, B.J. *et al.* (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485, 246–250
76 Neale, B.M. *et al.* (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485, 242–245
77 Morrow, E.M. *et al.* (2008) Identifying autism loci and genes by tracing recent shared ancestry. *Science* 321, 218–223
78 Gamsiz, E.D. *et al.* (2013) Intellectual disability is associated with increased runs of homozygosity in simplex autism. *Am. J. Hum. Genet.* 93, 103–109
79 Yu, T.W. *et al.* (2013) Using whole-exome sequencing to identify inherited causes of autism. *Neuron* 77, 259–273
80 Chahrour, M.H. *et al.* (2012) Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* 8, e1002635
81 Lim, E.T. *et al.* (2013) Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron* 77, 235–242
82 Mansour, H.A. *et al.* (2009) Association study of 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar Disord.* 11, 701–710
83 Mansour, H. *et al.* (2010) Consanguinity and increased risk for schizophrenia in Egypt. *Schizophr. Res.* 120, 108–112
84 Dratva, J. *et al.* (2007) Variability of reproductive history across the Swiss SAPALDIA cohort – patterns and main determinants. *Ann. Hum. Biol.* 34, 437–453
85 Ten Kate, L.P. *et al.* (2010) Autosomal recessive disease in children of consanguineous parents: inferences from the proportion of compound heterozygotes. *J. Community Genet.* 1, 37–40
86 Rauch, A. *et al.* (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380, 1674–1682
87 Ellison, J.W. *et al.* (2013) Genetic basis of intellectual disability. *Annu. Rev. Med.* 64, 441–450
88 Hoodfar, E. and Teebi, A.S. (1996) Genetic referrals of Middle Eastern origin in a western city: inbreeding and disease profile. *J. Med. Genet.* 33, 212–215
89 Hamamy, H.A. *et al.* (2007) Consanguinity and genetic disorders. Profile from Jordan. *Saudi Med. J.* 28, 1015–1017
90 Zoghbi, H.Y. and Warren, S.T. (2010) Neurogenetics: advancing the 'next-generation' of brain research. *Neuron* 68, 165–173
91 Lynch, M. (2010) Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 961–968
92 Vissers, L.E. *et al.* (2010) A de novo paradigm for mental retardation. *Nat. Genet.* 42, 1109–1112
93 Dixon-Salazar, T.J. *et al.* (2012) Exome sequencing can improve diagnosis and alter patient management. *Sci. Transl. Med.* 4, 138ra178
94 Bell, C.J. *et al.* (2011) Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci. Transl. Med.* 3, 65ra64
95 Kingsmore, S. (2012) Comprehensive carrier screening and molecular diagnostic testing for recessive childhood diseases. *PLoS Curr.* 1, 1–23 http://dx.doi.org/10.1371/e4f9877ab8ffa9 [Edition1]
96 Bittles, A.H. (2008) A community genetics perspective on consanguineous marriage. *Community Genet.* 11, 324–330
97 Modell, B. and Darr, A. (2002) Science and society: genetic counselling and customary consanguineous marriage. *Nat. Rev. Genet.* 3, 225–229
98 Wright, S. (1922) Coefficients of inbreeding and relationship. *Am. Nat.* 56, 330–338
99 Molinari, F. *et al.* (2002) Truncating neurotrypsin mutation in autosomal recessive nonsyndromic mental retardation. *Science* 298, 1779–1781
100 Higgins, J.J. *et al.* (2004) A mutation in a novel ATP-dependent Lon protease gene in a kindred with mild mental retardation. *Neurology* 63, 1927–1931
101 Basel-Vanagaite, L. *et al.* (2006) The CC2D1A, a member of a new gene family with C2 domains, is involved in autosomal recessive non-syndromic mental retardation. *J. Med. Genet.* 43, 203–210
102 Motazacker, M.M. *et al.* (2007) A defect in the ionotropic glutamate receptor 6 gene (GRIK2) is associated with autosomal recessive mental retardation. *Am. J. Hum. Genet.* 81, 792–798
103 Garshasbi, M. *et al.* (2008) A defect in the TUSC3 gene is associated with autosomal recessive mental retardation. *Am. J. Hum. Genet.* 82, 1158–1164
104 Molinari, F. *et al.* (2008) Oligosaccharyltransferase-subunit mutations in nonsyndromic mental retardation. *Am. J. Hum. Genet.* 82, 1150–1157
105 Mir, A. *et al.* (2009) Identification of mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation. *Am. J. Hum. Genet.* 85, 909–915
106 Philippe, O. *et al.* (2009) Combination of linkage mapping and microarray-expression analysis identifies NF-kappaB signaling defect as a cause of autosomal-recessive mental retardation. *Am. J. Hum. Genet.* 85, 903–908
107 Mochida, G.H. *et al.* (2009) A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. *Am. J. Hum. Genet.* 85, 897–902
108 Ratan, A. *et al.* (2013) Comparison of sequencing platforms for single nucleotide variant calls in a human sample. *PLoS ONE* 8, e55089
109 Ropers, F. *et al.* (2011) Identification of a novel candidate gene for non-syndromic autosomal recessive intellectual disability: the WASH complex member SWIP. *Hum. Mol. Genet.* 20, 2585–2590