APPLICATIONS OF NEXT-GENERATION SEQUENCING

Genetic studies in intellectual disability and related disorders

Lisenka E. L. M. Vissers[1], Christian Gilissen[1] and Joris A. Veltman[1,2]

Abstract | Genetic factors play a major part in intellectual disability (ID), but genetic studies have been complicated for a long time by the extreme clinical and genetic heterogeneity. Recently, progress has been made using different next-generation sequencing approaches in combination with new functional readout systems. This approach has provided novel insights into the biological pathways underlying ID, improved the diagnostic process and offered new targets for therapy. In this Review, we highlight the insights obtained from recent studies on the role of genetics in ID and its impact on diagnosis, prognosis and therapy. We also discuss the future directions of genetics research for ID and related neurodevelopmental disorders.

Epilepsy
Group of neurological diseases that are characterized by episodes of shaking, which can range in severity from brief, nearly undetectable to long and vigorous.

Autism spectrum disorder
(ASD). Collective term to describe a wide range of conditions that are characterized by social deficits and communication difficulties, stereotyped or repetitive behaviour and interests, sensory issues and, in some cases, cognitive delays.

[1]*Department of Human Genetics, Donders Centre for Neuroscience, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA Nijmegen, The Netherlands.*
[2]*Department of Clinical Genetics, Maastricht University Medical Centre. Universiteitssingel 50, 6229 ER Maastricht, The Netherlands.*
Correspondence to J.A.V.
joris.veltman@radboudumc.nl

doi:10.1038/nrg3999
Published online 27 Oct 2015

Intellectual disability (ID), the world-wide prevalence of which has been estimated at 1%[1], is characterized by substantial limitations in both intellectual functioning and adaptive behaviour, starting before the age of 18 years. Most individuals with ID are identified early in childhood because of developmental delays, and ID is a prominent feature of most developmental disorders. However, a formal diagnosis of ID is made only when IQ testing identifies an IQ score of less than 70. ID can occur in isolation or in combination with congenital malformations or other neurological features such as epilepsy, sensory impairment and autism spectrum disorders (ASD), and its severity (mild, moderate, severe and profound) is highly variable.

ID can be caused by exogenous factors such as maternal alcohol abuse during pregnancy, infections, birth complications and extreme malnutrition, but genetics is known to have an important role in its aetiology. The brain is an incredibly complex organ consisting of a myriad of interconnected cell types. During development and day-to-day functioning throughout life, numerous proteins need to be functionally active in the right amount at the right place and the correct time. It is therefore not unexpected that a mutation, deletion or rearrangement affecting any one of the genes encoding these proteins can have severe consequences for brain development or cognitive functioning. Indeed, family and population studies of intelligence show a high heritability, but no reliable literature exists on the heritability of ID itself. ID has become the most frequent reason for referral to paediatric genetic services. Unfortunately, the clinical heterogeneity of ID is reflected with extreme genetic heterogeneity, and a genetic diagnosis is still lacking in most cases[2]. This is concerning, as a genetic diagnosis can provide detailed information on the subtype of ID, its prognosis, possible complications, treatment options as well as information on the inheritance relevant for family planning. Genetic research has focused on the use of unbiased genome-wide approaches because of the genetic heterogeneity of ID. These include genomic microarrays and, more recently, next-generation sequencing (NGS) using extensive gene panels, the exome or the whole genome. With the introduction of NGS technologies, new ID genes are now being identified in rapid succession. This combination of novel technology and increased biological understanding is rapidly increasing the diagnostic yield of genetic tests in ID and improving the usefulness of genetic test results for patients and families involved. In addition, it is providing possibilities for carrier testing and prenatal screening, as well as new targets for treatment. In studying the genetics of ID, it is important to note that ID, which by definition is an early-onset disorder, has a significant impact on fitness. The exact effect on fitness will depend on the severity of the ID and the presence or absence of additional clinical features. For this reason, severe ID is mostly sporadic, whereas milder forms of ID can occur in families and spread through the population. This difference in fitness effects will impact the genetic architecture underlying these different forms of ID. Indeed, genomic microarray studies have revealed a strong relationship between the number of genes affected by rare copy number variants (CNVs) and the severity of the ID[3]. In addition, dominantly acting rare *de novo* mutations have recently

been shown to be a major cause of severe ID and associated developmental disorders[4–10]. By contrast, more common and complex forms of inheritance are expected to underlie the milder forms of ID. At present, however, much less is known about the genetics underlying these mild forms of ID. The genetics of ID also differs markedly in countries with a high frequency of consanguinity, owing to a more prominent role for recessive inheritance[11] (see REF. 12 for a recent review).

In this Review, we highlight the insights obtained from recent studies on the role of genetics in ID and its impact on diagnosis, prognosis and therapy. We mainly focus on the genetic aspects of ID and do not extensively discuss the clinical aspects of ID (for coverage of this topic, see REF. 13). First, we provide a brief historical overview of genomic approaches developed to study ID. Next, we discuss results from recent genomic studies, with a focus on both dominant *de novo* mutations, as well recessive forms of inheritance. In addition, we describe overlap with related disorders, such as ASD and epilepsy, as well as ID gene networks and pathways in which linked genes act in concert. Finally, we outline the future of ID research and emerging approaches for therapy.

A brief historical overview

Historically, genetic diagnosis of ID started under the microscope with the identification of trisomy 21 as the cause of Down syndrome[14] and a marker chromosome X for fragile X syndrome[15]. Because of these discoveries and the widespread introduction of cytogenetic banding technologies, chromosomal abnormalities were soon recognized as a common cause of ID, explaining up to 15% of cases[16,17], and Down syndrome was recognized as the most frequent genetic form of ID. In 1991, expansion of a CGG repeat in the fragile X mental retardation 1 (*FMR1*) gene was identified as the cause of fragile X syndrome[18,19], with an estimated frequency of 1 in 5,000 males and accounting for ~0.5% of ID[20]. *FMR1* remains the most mutated gene in ID and is routinely tested in ID diagnostics.

Chromosome X became a main focus of research in the 1990s and the beginning of this century because of the observed unbalanced sex ratio of 1.3–1.4 to 1 for male to females with ID and the possibility to carry out linkage analysis in large families with only male patients[21]. Mutations causing monogenic forms of X-linked ID have now been identified in more than 100 genes. None of these genes individually explains more than 0.1% of ID, but collectively they explain up to 10% of ID in males[22].

Research on autosomal causes of ID was initially hampered by the size of our genome, the limitations of cytogenetic and sequencing technologies and the absence of large families with autosomal forms of ID (owing to the effect of ID on reproductive fitness). Cytogenetic studies did reveal a number of recurrent autosomal microdeletions and duplications that explain clinically distinct ID syndromes such as Prader–Willi syndrome, Angelman syndrome, Williams syndrome, Smith–Magenis syndrome, Miller–Dieker syndrome and DiGeorge syndrome[23]. The introduction of genomic microarrays allowed the genome-wide detection of CNVs at finer resolution than was possible using the light microscope[24]. Pioneering microarray studies in ID showed that these CNVs occurred *de novo* in the germline of ~10% of patients[25–28]. Although these CNVs occurred all over the genome, numerous recurrent *de novo* CNVs have now been identified in ID as autosomal-dominant causes[23]. Detailed clinical and genetic characterization of these patients has resulted in the description of many new ID syndromes, provided insights into the genomic architecture underlying these genomic disorders and revealed many causative dosage-sensitive genes located in these CNV regions[29–31]. Because of their superior resolution and diagnostic yield, genomic microarrays rapidly replaced G-banded karyotyping as the first-tier test for ID[32]. In spite of these successes, the number of autosomal-dominant ID genes is still small relative to the number of X-linked ID genes.

In the meantime, over 300 genes have been identified for autosomal-recessive forms of ID, mostly by homozygosity mapping using single nucleotide polymorphism (SNP) microarrays and subsequent follow-up of candidate genes by Sanger sequencing. This number may be slightly inflated, as more than 97% of these genes have a role in recessive disorders that include ID as one of their main features. In fact, very few recessively inherited genes have been identified that only cause an ID phenotype (known as isolated ID)[12,33].

Overall, over 700 genes have now been identified across studies of X-linked, autosomal-dominant and autosomal-recessive ID, which can be used for the molecular diagnosis of both isolated ID and ID-associated disorders (FIG. 1; see Supplementary information S1 (table)).

From genes to genomes

NGS technologies have facilitated genetic research on ID in the past 5 years. The power of NGS approaches for disease gene identification was first shown in unexplained rare syndromes. Exome sequencing of four patients with Miller syndrome, a rare syndrome presumed to be recessively inherited, allowed the identification of dihydroorotate dehydrogenase (quinone) (*DHODH*) as the causative gene[34]. Exome sequencing also rapidly identified disease-causing genes for autosomal-dominant sporadic syndromes associated with ID such as Schinzel–Giedion syndrome[35], Kabuki syndrome[36] and Bohring–Opitz syndrome[37]. These studies showed that it is possible to prioritize disease-causing single nucleotide variants that are present in the exomes of a few patients with overlapping clinically defined syndromes. This unbiased NGS approach has accelerated disease gene identification for rare monogenic ID syndromes that could not have been effectively studied before the NGS era because of the inability to study all coding mutations in individual patients[38]. As a next step, exome sequencing was also applied to common, non-syndromic forms of ID.

Clinical heterogeneity
The phenomenon by which the same (genetic) disease can have differences in clinical manifestation.

Genetic heterogeneity
The phenomenon by which mutations in different genes can cause a similar phenotype.

Next-generation sequencing
(NGS). A collective term to describe the modern high-throughput sequencing technologies in the post-Sanger sequencing era.

Diagnostic yield
The percentage of patients who receive a conclusive molecular diagnosis for their disease.

Copy number variants
(CNVs). Insertions or deletions larger than 1,000 nucleotides in size.

De novo mutations
Genetic alterations that are present for the first time in one family member as a result of a mutation in the germ cell of one of the parents, or in the fertilized egg itself.

Down syndrome
The first recognized and most common form of a human aneuploidy syndrome, consisting of trisomy of chromosome 21 and representing 6–8% of all intellectual disability (ID) cases. The syndrome is generally associated with physical growth delays, characteristic facial features and mild to moderate ID.

Fragile X syndrome
Genetic disorder characterized by intellectual disability (ID), elongated face, large or protruding ears, macroorchidism, stereotypic movements and social anxiety.

Prader–Willi syndrome
Genetic condition that is characterized by hypotonia, feeding difficulties, poor growth, delayed development and behavioural problems. In infancy, patients develop an insatiable appetite that leads to chronic hyperphagia and obesity.

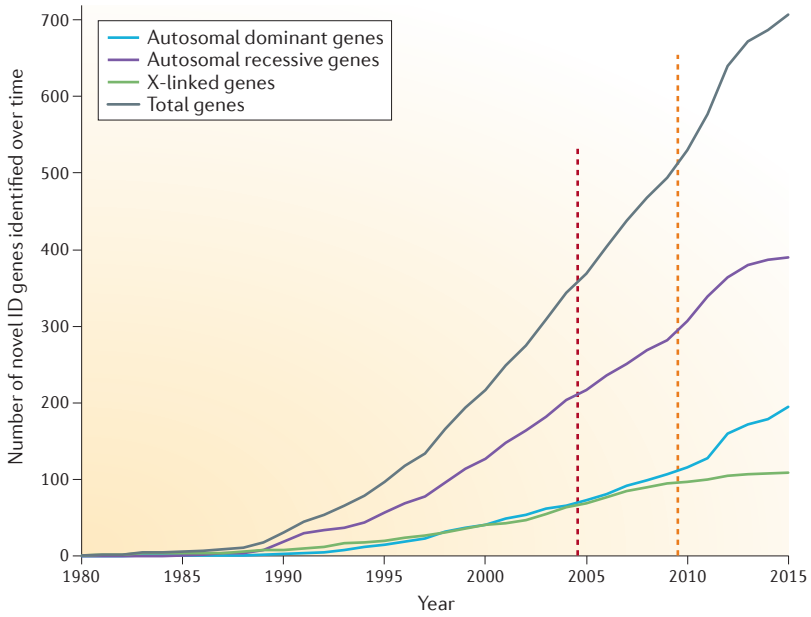

Figure 1 | **Increase of genes linked to isolated ID and ID-associated disorders.** Graphical overview of the increase in gene discovery for isolated intellectual disability (ID) and ID-associated disorders over time, specified by the type of inheritance. Vertical dashed lines represent the introduction of genomic microarrays (red) and next-generation sequencing (NGS)-based technologies (orange) for the detection of new ID genes. From this figure it is clear that we have not reached any saturation in ID disease gene identification, except perhaps for X-linked forms of ID. Supplementary information S1 (table) lists all genes shown in this figure, along with their respective ID phenotype.

Angelman syndrome
Complex genetic disorder that primarily affects the nervous system and is characterized by delayed development, intellectual disability, speech impairment, epilepsy and problems with movement and balance. Patients typically have a happy, excitable, demeanour.

Williams syndrome
Genetic condition that is characterized by mild to moderate intellectual disability, cardiovascular disease, distinctive facial features and a typical outgoing, engaging personality.

Smith–Magenis syndrome
Developmental disorder that is characterized by mild to moderate intellectual disability, delayed speech and language skills, distinctive facial features, sleep disturbances and behavioural problems.

A de novo paradigm in ID. The genetic heterogeneity observed in common forms of ID reduces the possibility to prioritize disease genes on the basis of overlapping mutations in multiple patients. Potentially pathogenic mutations should therefore be identified in individual patients. Vissers *et al.*[8] approached this for the first time by sequencing the exomes of ten patients with unexplained severe ID as well as their unaffected parents. This trio-based exome sequencing design allowed the identification of *de novo* dominant mutations as a possible cause of sporadic forms of ID. *De novo* mutations had been hypothesized to represent a common cause of severe ID before these exome sequencing studies, potentially explaining why this severe early-onset disorder remains so frequent in our population[39]. In line with this hypothesis, and as described above, *de novo* CNVs were already known to represent the most common cause of ID in the Western world. In addition, Sanger sequencing of candidate ID genes, although laborious when performed in large numbers of cases, had been applied successfully to reveal *de novo* mutations in newly described ID genes such as synaptic RAS GTPase activating protein 1 (*SYNGAP1*)[40] and forkhead box P1 (*FOXP1*)[41]. In the pilot exome sequencing study by Vissers *et al.*[8], *de novo* mutations predicted to be damaging were identified in two known ID genes as well as in four candidate ID genes, potentially providing a genetic cause in six out of ten patients. Trio-based exome sequencing rapidly became a useful method for the unbiased identification of *de novo* mutations in ID and other early-onset neurodevelopmental disorders[42–49].

In 2012, two studies confirmed the importance of *de novo* mutations in larger cohorts of 50 to 100 patients with unexplained severe ID, and demonstrated the utility of exome sequencing as a diagnostic test[4,7]. These studies indicated that *de novo* mutations with a predicted damaging effect on currently known ID genes can explain 13–35% of cases with severe ID. A similar result was recently obtained in a study of 41 patients with moderate to severe ID[50], with a diagnostic yield of ~29% based only on *de novo* mutations in known ID genes. Also in line with this, a much larger trio-based exome sequencing study in 1,133 children with severe undiagnosed developmental disorders (87% with ID or developmental delay) reported a diagnostic yield of 18% for *de novo* mutations in known disease genes[10]. Differences in the diagnostic yield between these studies can be explained by differences in patient selection, in variant interpretation and in the quality of the exome sequencing approach.

Recently, a follow-up study[5] was carried out using whole-genome sequencing in 50 of the 100 patients included in one of the exome sequencing studies mentioned above[4]. Genome sequencing identified 84 *de novo* coding mutations in these patients. Of note, 65 of these coding mutations (77%) were missed in the original exome sequencing study owing to poor coverage. Likewise, whole-genome sequencing detected many *de novo* genomic structural variants missed by previous microarray studies in these patients. These variants included two *de novo* single exon deletions and one intra-exonic deletion, both affecting known ID genes. Based on this first whole-genome sequencing study in severe ID[5], the authors estimated that 60% of severe ID can be explained by a *de novo* coding mutation in a known ID gene (39% by *de novo* point mutations and 21% by *de novo* structural variants; see FIG. 2 for an overview of the recent increase in diagnostic yield for ID). Even though whole-genome sequencing was used in this study, no evidently pathogenic non-coding mutations were identified, exemplifying that variant interpretation outside of genes is still highly complex (see the 'New research directions' section below).

Next-generation recessive disease gene identification. Although recessive inheritance may not be a major cause of ID in outbred populations, it is perceived as a key cause in inbred populations[12]. In 2011, targeted NGS was used systematically for the first time after homozygosity mapping, and revealed homozygous regions in 136 consanguineous families with autosomal-recessive ID[11]. This study identified recessively inherited mutations in 23 known ID genes and 50 new candidate ID genes. Five of these genes were mutated in more than one family (three known ID genes and two candidate ID genes, each mutated in two families). In total, probable causative homozygous mutations were identified in 78 of the 136 families, potentially explaining 57% of ID cases.

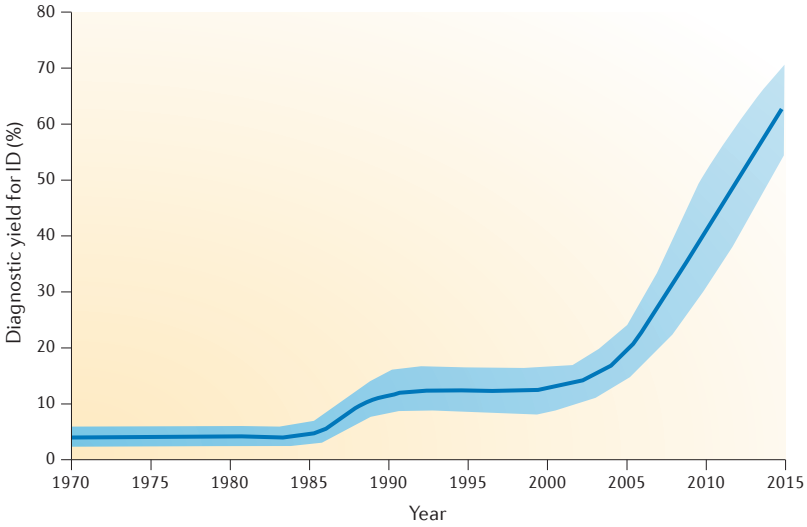

Figure 2 | **Diagnostic yield for ID over time.** Graphical overview of the diagnostic yield for moderate to severe intellectual disability (ID) (excluding Down syndrome, which represents 6–8% of all ID) over time. Solid line indicates the mean of published studies, and the shaded background indicates the lower and upper boundaries of reported diagnostic yields. In the 1970s, conventional karyotyping became a routine diagnostic test and provided a conclusive diagnosis in 3–6.5% of ID cases. The diagnostic yield increased by 6–10% after the introduction of both Sanger sequencing and targeted fluorescence *in situ* hybridization (FISH) in the 1990s[120]. At the beginning of this century, genomic microarrays were introduced, increasing the diagnostic yield by another 15–23%[25,32]. The introduction of whole-exome sequencing in 2010 and onwards added a diagnostic yield of 24–33%[4,7,10], and a first pilot study using whole-genome sequencing added a further 26% in 2014 (REF. 5), accumulating to an overall diagnostic yield of 55–70% for moderate to severe ID. Interestingly, a higher diagnostic yield has been observed for moderate to severe ID (IQ score <50) compared with mild ID (IQ score 50–70)[120–122]. As an example, subtelomeric aberrations explain 0.5% of mild ID and 7.4% of moderate to severe ID. Since the introduction of genomic microarray technology, the diagnostic yield per category of ID is less well documented. This is also the case for differences in diagnostic yield between males and females.

Miller–Dieker syndrome
Genetic condition that is characterized by a pattern of abnormal brain development known as lissencephaly, which leads to severe intellectual disability, developmental delay, seizures, spasticity, hypotonia and feeding difficulties.

DiGeorge syndrome
Genetic disorder that is characterized by a heart defect, learning difficulties and cleft palate, among other symptoms.

G-banded karyotyping
Visualization of the chromosome count present in the nucleus of a eukaryotic cell after Giemsa staining, followed by trypsin digestion, using a light microscope. The staining results in a recognizable pattern of light (euchromatic) and dark (heterochromatic) stained bands.

In recent years, exome sequencing has also become the standard technology to identify recessive causes of ID. This has resulted in the description of numerous recessive ID genes such as DDHD domain containing 2 (*DDHD2*), methyltransferase like 23 (*METTL23*), calpain 10 (*CAPN10*), serine/threonine/tyrosine interacting-like 1 (*STYXL1*) and solute carrier family 6 (neutral amino acid transporter) member 17 (*SLC6A17*)[51–53].

From candidate to validated ID genes. The focus on dominant *de novo* or recessively inherited mutations that are present in exome and genome sequencing data allows one to effectively prioritize disease-causing mutations and to identify new candidate ID genes. However, proving pathogenicity of these mutations and establishing that a candidate ID gene is an ID-causing gene when mutated is still complicated, even in this post-NGS era (see REF. 54 for guidelines on investigating the causality of rare mutations in disease, and REFS 55,56 for reviews on the interpretation of *de novo* mutations).

Replication — the observation of multiple unrelated but phenotypically similar patients with mutations that are predicted to be damaging in the same gene — is still a key requirement in this process. The combination of CNV morbidity maps with lists of candidate ID genes that are affected by point mutations can be particularly powerful in this respect because both have an important role in ID[57]. In addition, an association with disease has to be statistically established by comparing the frequency of these types of mutations in cases and controls, especially as it is now well established that mutation rates vary substantially between genes. Alternatively, for *de novo* mutations, one can use estimated gene-specific mutation rates to demonstrate an enrichment of gene mutations within a patient cohort[56]. Of note, a recent study re-assessed published X-linked ID-causing mutations in 106 genes by analysing the frequency of truncating variants of these genes in controls using the US National Heart, Lung and Blood Institute (NHLBI) large-scale sequencing database. The authors questioned whether the mutations in ten of these genes cause X-linked ID, as truncating mutations are also present in controls. For 15 additional genes, further studies were required to conclusively implicate them as causes of X-linked ID[58].

The replication and association of new candidate ID genes identified through exome or whole-genome sequencing can be achieved by targeted re-sequencing using high-throughput NGS approaches in thousands of cases and controls[57,59]. Within the context of the Autism Spectrum/Intellectual Disability network (ASID), more than 15,000 patients and control samples have been collected for replication studies[60]. Importantly, detailed clinical follow-up studies and analysis of parental samples are required for all patients with recurrently mutated genes to establish genotype–phenotype correlations and the mode of inheritance. This reverse phenotyping approach, in which genetic data are used to drive new phenotypic definitions, has already revealed a number of new ID genes and allows the analysis of the role of these genes in other neurodevelopmental disorders[61,62].

Complementary to finding the recurrence of specific mutated genes in patients, and the absence in controls, the true nature of candidate ID genes and of their mutations must be established by carrying out functional studies, either *in vitro* using patient-derived cells, or *in vivo* using animal models. Mice have historically been used as model organisms for ID to learn about disease biology and to find potential therapeutic strategies for specific diseases, for instance for fragile X syndrome[63]. More recently, fruit flies and zebrafish have been introduced as disease models for ID, largely because of their reduced costs and short generation times[64,65]. Despite the evolutionary distances, 73% of human ID genes have been reported to have an unambiguous counterpart in fruit flies[66], and zebrafish have 70% overall genetic identity to humans[64]. The fruit fly offers the opportunity to study behaviour using various paradigms including olfactory learning and courtship conditioning, spatial learning and habituation assays. In addition, the use of fruit flies allows the study of brain organization and architecture during neuronal development (reviewed in REF. 67). Also in zebrafish, a large variety of assays is available to monitor development (reviewed in REF. 64).

Homozygosity mapping
A method to map human recessive disease traits with DNA of inbred children.

Schinzel–Giedion syndrome
A rare genetic disorder of congenital hydronephrosis, skeletal dysplasia and severe developmental retardation.

Kabuki syndrome
A rare genetic condition that is characterized by distinctive facial features, skeletal anomalies and intellectual disability.

Although model animals are essential in ID research, it would be ideal to study the function of normal and mutated ID genes in humans. This is complicated by the fact that the affected tissue, the brain, cannot be easily accessed. The possibility to make patient-derived induced pluripotent stem cells (iPSCs) will undoubtedly accelerate functional studies for ID[68]. Together with the recent introduction of CRISPR–Cas9 gene editing technology, it has now become possible to model disease-specific mutations, rather than overall gene knockdown, in both model organisms and human cells[69]. This will allow a much more accurate characterization of the consequences of mutations on brain phenotypes, even if these mutations are observed in a single patient and/or family.

Towards a comprehensive set of ID genes. As mentioned above, ~700 genes have been convincingly linked to either isolated, non-syndromic forms of ID, or to a disorder that includes ID as one of its major features (FIG. 1; see Supplementary information S1 (table)). On the basis of whether the plots in FIG. 1 are reaching a plateau, it seems that most X-linked ID genes have been identified by now, whereas many (if not most) autosomal ID genes still await discovery. It is difficult to reliably predict how many ID genes are still to be discovered and when all ID genes will be identified (see for example REF. 70 for an estimation of the number of autism genes), but the number is likely to exceed 1,000 in the coming decade. One could expect that most of the frequently mutated genes should have been identified by now, but even this may not be the case. As an illustration, two recently discovered ID genes, AT-rich interactive domain 1B (SWI1-like) (*ARID1B*) and DEAD (Asp-Glu-Ala-Asp) box helicase 3, X-linked (*DDX3X*) may each explain more than 1% of ID patients[71,72].

Biology underlying ID and related disorders

Biological processes affected in ID. The proteins encoded by ID genes can be hypothesized to play a part in one or more shared pathways or functional modules, either through direct interactions or as part of more complex interaction networks. Knowledge of the affected pathways and modules may form important targets for therapy development[73,74] (BOX 1). Commonly used approaches to identify such genetic networks use enrichment analysis based on gene ontology terms. This has revealed overt cellular processes including neurogenesis, neuronal migration, synaptic function and regulation of transcription and translation[74]. However, translating these biological processes into common genetic networks is more challenging as there is often a way to link genes and proteins to one another, which can lead to presumptuous conclusions of networks involved. Such incorrect assumptions are mostly due to three reasons: the function of most known genes is not fully understood; the grouping of affected genes is often arbitrary; and the concepts of pathways and networks are based on biochemistry, which may not be appropriately capturing the complex scenarios of the true biological system[75].

Despite these limitations, biological processes involved in ID are starting to emerge. For a long time, cellular signalling pathways have been known to have an important role in the aetiology of ID. As an example, the RAS–MAPK (mitogen-activating protein kinase) pathway is associated with a particular set of intellectual disabilities, the so-called rasopathies. These encompass well-known ID syndromes, such as Noonan syndrome and Costello syndrome. Mutations in this pathway impede the correct functioning of the MAPK signalling cascade, a metabolic pathway that regulates growth factors and embryological development[76].

Another emerging cellular signalling cascade is the RHO GTPase pathway, which consists of guanine-nucleotide binding proteins that act as 'molecular switches' in a wide variety of cellular functions, including the morphogenesis of dendritic spines, which are crucial for learning and memory. In total, over 20 GTPases are known, of which the effectors RAC1, cell division cycle 42 (CDC42) and RHOA have established roles in spine formation and synapse plasticity. Mutations in both regulators and effectors

Box 1 | Therapy for intellectual disability

For many years, the prevailing view has been that intellectual disability (ID) cannot be cured because the affected biological processes, including neurogenesis, are difficult to target and the cellular damage cannot be undone or reversed. Much of the treatment has therefore focused on environmental optimization, including individualized education plans, as well as minimizing complicating co-morbidities (for example, visual, sleep or pain co-morbidities)[112]. For specific syndromes associated with ID (see REF. 112 for a review), some therapeutic strategies are known. For instance, for a few metabolic disorders, enzyme replacement therapy is used, which can drastically change prognosis, and is sometimes accompanied by unexpected intellectual sparing (for example, for Pompe disease).

However, recent studies involving specific types of ID, including for instance those resulting from genes involved in histone modification, suggest that ID might actually be amenable to therapeutic intervention. For example, Kabuki syndrome is caused by *de novo* mutations in lysine (K)-specific methyltransferase 2D (*KMT2D*) or lysine (K)-specific demethylase 6A (*KDM6A*)[36,113], and probably results from an imbalance in open and closed chromatin states. Recently, it has been shown that memory deficits in the mouse model can be prevented, or even reversed, through systemic delivery of drugs that directly influence the histone modification events that favour chromatin opening[114]. The authors of the study speculate that other genetic disorders involving the histone modification machinery might also be amenable to therapeutic intervention with histone deacetylase (HDAC) inhibitors[115].

In keeping with this concept, it has long been hypothesized that neurodevelopmental disorders caused by genes in the same network may be targeted by a central node in the system, and it has been argued that the GABAergic system — which is disrupted in fragile X syndrome, Rett syndrome, methyl CpG binding protein 2 (*MECP2*) duplication syndrome and Dravet syndrome — is a key candidate to target for therapeutic intervention[116]. Indeed, mouse studies on *MECP2* duplication syndrome have shown that chronic treatment with low doses of $GABA_A$ receptor antagonists ameliorates specific behavioural phenotypes, including motor coordination, episodic memory impairments and synaptic plasticity deficits[117].

Inspired by these studies in model organisms, clinical trials interfering with the GABAergic system have been set up (reviewed in REF. 116). Interestingly, early intervention combined with targeted treatment in young children with fragile X syndrome has been shown to improve behaviour and cognition, suggesting that targeted pharmacological treatments have great potential[118]. In general, the road ahead for treatment of ID will mostly have to focus on the identification of the commonly disturbed pathways and networks, as ID is too heterogeneous for treatment at the individual gene level. With the advent of stem cell replacement, future research will also reveal to what extent cellular replacement for ID and other neurodevelopmental disorders resulting from reduced numbers of specific neurons may prove rational[119].

of the RHO GTPases have been found to underlie various forms of non-syndromic ID[77]. In addition, mutations in the downstream effectors of the RHO GTPases, including phosphatases and calcium/calmodulin-dependent protein kinase type II (CaMKII) subunits, have been reported in patients with ID[4,78].

Transcription regulation and chromatin remodelling are deregulated by mutations in euchromatic histone-lysine *N*-methyltransferase 1 (*EHMT1*), leading to Kleefstra syndrome[79], and by mutations in the SWI/SNF chromatin remodelling complex, as observed in Coffin–Siris syndrome[72]. The involvement of these biological processes in ID also hints towards an overlap with carcinogenesis, and the timing of the *de novo* mutation in these pathways seems to determine the clinical outcome. For example, *de novo* germline mutation p.Arg1075* in *ARID1B* leads to Coffin–Siris syndrome, whereas the same mutation occurring later in life can lead to cancer[72,80]. Similarly, *de novo* germline mutations in the degron consensus sequence of SET binding protein 1 (*SETBP1*), which lead to Schinzel–Giedion syndrome[35], are also observed as somatic mutations in myeloid cancers [80].

Overlap with other neurodevelopmental disorders. Many studies have suggested shared molecular pathways for ID and other neurodevelopmental disorders[81,82]. This has been inspired by the high co-morbidity that is commonly observed between ID and other cognitive impairments, such as autism, schizophrenia and epilepsy[74].

Recent large-scale patient–parent trio sequencing studies have provided statistical evidence for an enrichment of *de novo* mutations in a small number of genes across different neurodevelopmental phenotypes[83]. Loss-of-function (LoF) *de novo* mutations identified in all of these studies show a pattern of significant overlap of the underlying genes (FIG. 3). In particular, ID and ASD show a large overlap: 17% of all genes with *de novo* LoF mutations observed in ID are also reported in ASD. Such an overlap in genes for ID and ASD has also been noted in other studies[70]. However, about 70% of individuals with autism also present with ID[84]. It is therefore unclear at present whether this molecular overlap points to shared biological pathways between two distinct neurodevelopmental disorders or whether it is explained by patients presenting with both phenotypes. Of interest in this respect, male patients with ASD harbouring *de novo* LoF mutations were found to have a lower non-verbal IQ than ASD males without *de novo* LoF mutations[85].

Mutations in some genes give rise to a wide variety of neurodevelopmental phenotypes. For example, mutations in sodium channel, voltage gated, type II α-subunit (*SCN2A*) have been identified in individuals with ID, ASD, epilepsy and ataxia[7,86,87]. In particular, genes involved in epigenetic regulation and chromatin remodelling, such as chromodomain helicase DNA binding protein 8 (*CHD8*), methyl CpG binding protein 2 (*MECP2*), autism susceptibility candidate 2 (*AUTS2*) and trafficking protein particle complex 9 (*TRAPPC9*), have been shown to have a role in a wide range of neurodevelopmental phenotypes[88]. The different manifestations of neurological phenotypes for mutations within the same gene are attributed to stochastic processes during development, the difference in genetic background between patients and the effects of different mutations. As an example, deletions in the gene hyperpolarization activated cyclic nucleotide gated potassium channel 1 (*HCN1*) cause a phenotype of combined ID and ASD, without epilepsy, whereas specific gain-of-function mutations in *HCN1* cause an isolated epilepsy phenotype without ID and/or ASD[89]. By contrast, mutations in other genes involved in potassium channel functioning have been identified exclusively in epilepsy-related phenotypes, sometimes with additional ID and/or ASD phenotypes[90]. This suggests that some basic molecular pathways are shared between all neurodevelopmental disorders in which defective genes will always give rise to multiple neurodevelopmental phenotypes, whereas others are more exclusively linked to a particular neurological disorder.

New research directions

In many ways genetic studies in ID are only just starting, with most research having focused so far on germline highly penetrant monogenic causes of ID. This research has been successful in explaining most severe forms of ID, but leaves a large proportion unexplained, especially for the milder forms of ID that lack additional malformations or neurological features. Genetic studies are now being expanded in at least three different but complementary directions that are described below: studying somatic causes of ID; studying digenic and oligogenic inheritance forms of ID; and studying non-coding causes of ID.

Somatic causes of ID. Somatic mutations, which accumulate in our cells during development and throughout life, can result in disease if they affect sufficient numbers of cells that develop and function differently because of these mutations (for a review, see REF. 91). For ID, it could well be that somatic mutations in neuronal cells (or their precursors) contribute to disease, but the question is how to identify these and study their impact on ID. Two of the main obstacles in studying somatic mutations are tissue sampling and the sensitivity of genomics technologies. Studies using peripheral blood DNA have recently shown that a considerable proportion of point mutations and CNVs occur post-zygotically in either the child or the parents[92,93]. These studies also indicate that highly sensitive genomics approaches are required to differentiate germline from somatic mutations and point to the importance of this distinction for estimating the recurrence risk in families. Related to this, a number of overgrowth and cortical malformation syndromes have recently been found to be caused by somatic mutations that are sometimes, but not always, detectable in DNA derived from blood[94–96].

These studies show the importance of analysing brain samples or even individual neurons to truly understand the role of somatic mutations in brain disorders such as ID. Single-cell sequencing, although still far from perfect because of amplification biases, has indicated that

Bohring–Opitz syndrome
A rare genetic disorder that is characterized by facial anomalies, multiple malformations, failure to thrive and severe intellectual disabilities.

Single nucleotide variants
(SNVs). Differences in the nucleotide composition at single positions in the genome.

Neurodevelopmental disorders
Term generally used to collectively describe disorders affecting neurodevelopment, including autism spectrum disorder, epilepsy, schizophrenia and intellectual disability.

Structural variants
Genomic regions of at least 1 kb in size that alter the normal chromosomal composition, such as inversions, translocations or copy number variants.

Induced pluripotent stem cells
(iPSCs). Adult cells that have been reprogrammed to stem cells and can thus be differentiated into different cell types.

Kleefstra syndrome
Genetic condition characterized by the core phenotype of developmental delay, intellectual disability, speech impairment, (childhood) hypotonia and distinct facial features, including synophrys, hypertelorism, midface hypoplasia, anteverted nares, prognathism, rolled out (everted) lips and macroglossia.

Coffin–Siris syndrome
A rare genetic disorder that causes developmental delays and absence of the fifth finger and toe nails.

Degron consensus sequence
Specific sequence of amino acids in a protein that directs the starting place of degradation.

Somatic mutations
Mutations that are present in a proportion of cells of the body except sperm and egg cells.

Schizophrenia
Psychotic disorder marked by severely impaired thinking, emotions and behaviour, including the inability to filter sensory stimuli and enhanced perceptions of sounds and colours.

Loss-of-function
(LoF). A mutation expected to result in reduced or abolished protein function.

somatic CNVs are present in most human neurons[97,98]. Some of these CNVs seem to have been generated during neurogenesis and are detected in multiple neurons, strengthening the hypothesis that somatic mutations in neurogenesis may contribute to disorders such as ID.

Increasing inheritance complexity in ID. Genotype–phenotype correlation studies in ID indicate that phenotypes are only rarely explained completely by a mutation in a single gene. Even in patients with the exact same *de novo* germline mutation, such as recently identified in the phosphofurin acidic cluster sorting protein 1 (*PACS1*) gene, clinical differences can be observed that are related to urogenital anomalies, cerebellum malformation, scoliosis and heart defects[99]. This phenotypic variation is not unexpected given that each individual inherits another 6 billion nucleotides in addition to this one mutation.

Variation in these nucleotides may affect the expressivity of disease as well as disease penetrance. Variable penetrance and expressivity has been noted for numerous well-defined CNVs associated with ID[3,100,101]. Systematic analysis of CNV data from thousands of genomic microarrays of patients with developmental delay has revealed substantial phenotypic variation among patients with identical rare CNVs[102]. This has prompted scientists to look for, and identify, additional rare CNVs in these

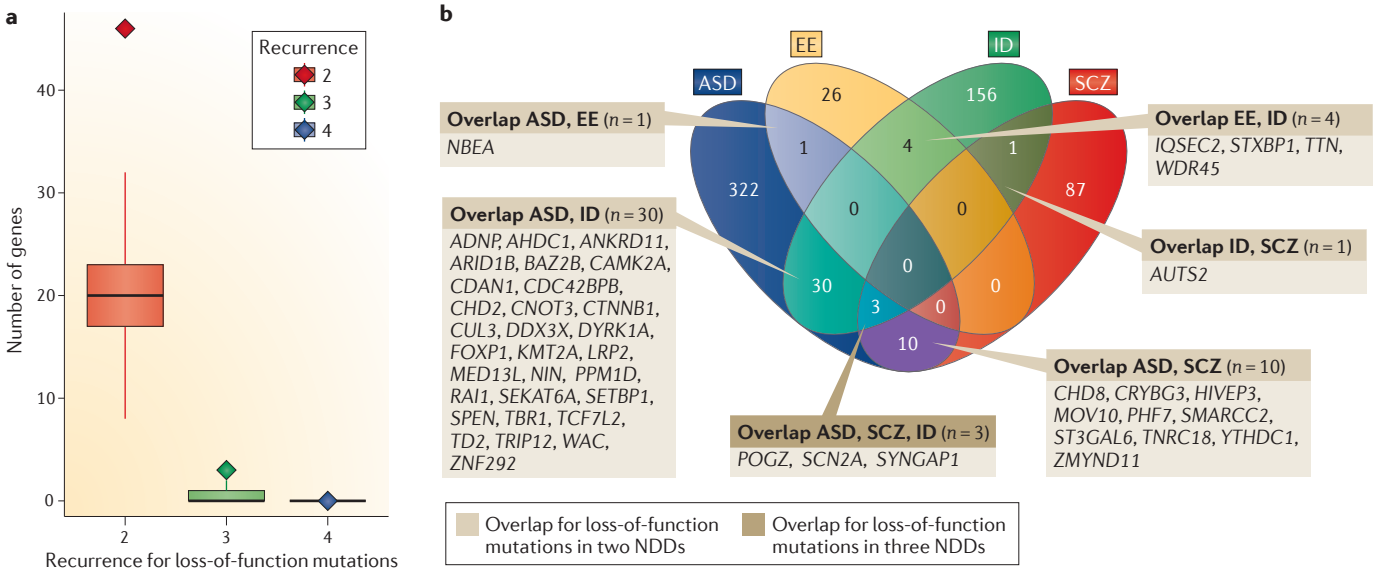

Figure 3 | **Genic overlap for neurodevelopmental disorders.** We collected *de novo* mutations of published patient–parent trio-based sequencing studies in neurodevelopmental disorders. All *de novo* mutations were re-annotated using our custom pipeline and grouped by phenotype: autism spectrum disorder (ASD; 2,683 patients)[47,116], epileptic encephalopathy (EE; 264 patients)[43], intellectual disability (ID; 1,284 patients)[4,5,7,10] and schizophrenia (SCZ; 1,063 patients)[44,46,49,89,117]. To assess the significance for overlap for *de novo* loss-of-function (LoF) mutations between these four neurodevelopmental disorders, we carried out 10,000 simulations with the total number of identified *de novo* mutations in these studies, making use of the gene-specific mutation rates from a previous study[56] **a** | The number of genes with overlapping *de novo* LoF mutations in two, three, and all four of the disorders, from 10,000 simulations, indicated as boxplots. Diamond symbols indicate the actual number of genes with *de novo* mutations across the neurodevelopmental disorders. There were significantly more genes with actual *de novo* LoF mutation for two and three disorders than expected by chance from the simulation studies ($P < 0.0001$ and $P = 0.0084$ respectively), whereas no genes with *de novo* LoF mutations in all four disorders were identified. **b** | Venn diagram denoting the overlap for the actual number of genes with *de novo* LoF mutations shared between each of the disorders. The genes for which overlap in *de novo* LoF mutations between neurodevelopmental disorders has been identified are listed. Importantly, this does not imply that all of these mutations are relevant for these neurodevelopmental disorders. This is because some of these genes, such as *TTN* (titin), have a high mutation rate and therefore *de novo* LoF mutations are also observed in unaffected individuals. *ADNP*, activity-dependent neuroprotector homeobox; *AHDC1*, AT hook, DNA-binding motif, containing 1; *ANKRD11*, ankyrin repeat domain 11; *ARID1B*, AT-rich interactive domain 1B (SWI1-like); *AUTS2*, autism susceptibility candidate 2; *BAZ2B*, bromodomain adjacent to zinc finger domain, 2B; *CAMK2A*, calcium/calmodulin-dependent protein kinase II-α; *CDAN1*, codanin 1; *CDC42BPB*, CDC42 binding protein kinase beta (DMPK-like); *CHD*, chromodomain helicase DNA binding protein; *CNOT3*, CCR4-NOT transcription complex, subunit 3; CRYBG3, βγ-crystallin domain-containing 3; *CTNNB1*, catenin (cadherin-associated protein), β1, 88kDa; *CUL3*, cullin 3; *DDX3X*, DEAD (Asp-Glu-Ala-Asp) box helicase 3, X-linked; *DYRK1A*, dual-specificity Tyr-phosphorylation regulated kinase 1A; *FOXP1*, forkhead box P1; *HIVEP3*, HIV type I enhancer binding protein 3; *IQSEC2*, IQ motif and Sec7 domain 2; *KMT2A*, lysine (K)-specific methyltransferase 2A; *LRP2*, low-density lipoprotein receptor-related protein 2; *MED13L*, mediator complex subunit 13-like; MOV10, Mov10 RISC complex RNA helicase; *NIN*, ninein (GSK3B interacting protein); *POGZ*, pogo transposable element with ZNF domain; *PPM1D*, protein phosphatase, Mg2+/Mn2+ dependent, 1D; *NBEA*, neurobeachin; *PHF7*, PHD finger protein 7; *RAI1*, retinoic acid induced 1; *SCN2A*; sodium channel, voltage gated, type II α-subunit; *SETBP1*, SET binding protein 1; *SMARCC2*, SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2; *ST3GAL6*, ST3 β-galactoside α-2,3-sialyltransferase 6; *STXBP1*, syntaxin binding protein 1; *SYNGAP1*, synaptic RAS GTPase activating protein 1; *TBR1*, T-box brain gene 1; *TCF7L2*, transcription factor 7-like 2 (T-cell specific, HMG-box); *TNRC18*, trinucleotide repeat containing 18; *TRIP12*, thyroid hormone receptor interactor 12; *WAC*, WW domain containing adaptor with coiled-coil; *WDR45*, WD repeat domain 45; *YTHDC1*, YTH domain containing 1; *ZMYND11*, zinc finger, MYND-type containing 11; *ZNF292*, zinc finger protein 292.

Digenic and oligogenic inheritance
A form of disease inheritance in which mutations in two (digenic) or more (oligogenic) unlinked genes must be present in a single individual to cause disease, whereas each mutation individually is insufficient to cause a phenotype.

Expressivity
The severity of the disease in individuals who have both the risk variant and the disease.

Penetrance
The proportion of patients with a specific phenotype among all carriers of a specific genotype.

Incidental findings
Medically relevant genetic variants unrelated to the clinical indication for which the genetic test was requested.

patients and their parents, which may partly explain the phenotypic heterogeneity. In addition, these studies have pointed to a sex bias in that male patients show more phenotypic variation and often inherit deleterious CNVs from their unaffected mother[3]. As a result, a 'female protective model' has been proposed, in which females require more severe mutations for neurodevelopmental disorders such as ID to develop[103].

So far, this type of complex inheritance has been studied only for the most rare and deleterious types of genomic variation — that is, large (>400 kb) *de novo* CNVs as well as *de novo* point mutations that affect neurodevelopmental genes. More comprehensive analysis of the role of complex forms of inheritance in ID will require much larger genome sequencing datasets of both patients and their family members. The challenges here are to have sufficient power to detect genetic interaction and to have sufficiently detailed and objective phenotypic information of patients and family members to do so.

Non-coding causes of ID. Although 98% of our genome is non-coding, no more than 3% of all known disease-causing mutations are located outside of genes[104]. Until recently it was impossible to reliably and affordably identify all non-coding variation in the genome and study potential phenotypic consequences. However, this is rapidly changing for the following three reasons: whole-genome sequencing will soon allow us to systematically detect all variation present in a patient's genome; new methods that have been developed and applied by members of the Encyclopedia of DNA Elements (ENCODE) consortium[105] and others[106] have identified gene-specific enhancer, insulator and repressor elements throughout the genome; and the CRISPR–Cas9 genome editing technology can be used to rapidly study the phenotypic consequences of mutations in these non-coding elements.

Although no breakthroughs using CRIPR–Cas9 genome editing have been published yet for ID, the power of this approach was recently demonstrated, when non-coding structural variants were linked to rare limb malformations[106]. In addition, recessive point mutations were recently detected in distal enhancers of pancreas-specific transcription factor 1A (*PTF1A*), a gene known to be involved in isolated pancreatic agenesis[107], and X-linked recessive point mutations were identified in a YY1-transcription binding site of the host cell factor C1 (*HCFC1*) gene, causing non-syndromic ID[108]. These initial successes show that it is possible to detect non-coding causes of disease. Most genomic variation, however, occurs in this part of the genome, and so the main challenge will be to prioritize variants for follow-up studies. For ID, a focus on *de novo* non-coding mutations that affect highly conserved promoters, transcription factor-binding sites, enhancers or insulator elements in the vicinity of known ID genes would be appropriate. In addition, *de novo* structural variants such as non-coding CNVs may be good targets for investigation as these are rare in the healthy population and are more likely to disrupt the regulatory landscape of the genome and result in disease[109].

Concluding remarks

Genetic studies have substantially improved our understanding of the causes of ID in the past decade, and for the first time we can provide a molecular diagnosis for most patients with severe ID. The genetic heterogeneity of ID requires genome-wide approaches, and genome sequencing is likely to become the first-tier diagnostic test for ID as soon as it is available and affordable. This will require appropriate counselling, and any incidental findings will need to be assessed without compromising a child's right to an open future[110]. It will take many more years before all ID-causing genes are identified, and the quest to identify and reliably interpret non-coding, somatic and complex genetic causes of ID is only just beginning.

However, even with incomplete information on ID-causing mutations it seems that *de novo* mutations are the main cause of severe ID. As a consequence, carrier screening to prevent transmission of inherited mutations to the offspring will not be a broadly useful strategy to prevent ID, at least not in outbred populations. Of note, most *de novo* mutations occur on the paternal allele and their number increases with paternal age[111]. Preventive strategies should therefore be focused on promoting early child bearing and/or freezing of sperm for future pregnancies. Finally, it is promising to see that therapeutic options are becoming a reality for a small subset of patients with genetic forms of ID. This demonstrates the importance of these genetic studies, now and in the future.

1. Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T. & Saxena, S. Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res. Dev. Disabil.* **32**, 419–436 (2011).
2. Mefford, H. C., Batshaw, M. L. & Hoffman, E. P. Genomics, intellectual disability, and autism. *N. Engl. J. Med.* **366**, 733–743 (2012).
3. Girirajan, S. *et al.* Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N. Engl. J. Med.* **367**, 1321–1331 (2012).
4. de Ligt, J. *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* **367**, 1921–1929 (2012).
5. Gilissen, C. *et al.* Genome sequencing identifies major causes of severe intellectual disability. *Nature* **511**, 344–347 (2014).
The first study to apply patient–parent trio-based whole-genome sequencing in a cohort of 50 patients with unexplained ID. A conclusive diagnosis was made in 21 of 50 patients. Also, it was estimated that genome sequencing as a first-tier test would identify the genetic cause in 62% of severe ID cases.
6. Hamdan, F. F. *et al.* Excess of *de novo* deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **88**, 306–316 (2011).
7. Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* **380**, 1674–1682 (2012).
8. Vissers, L. E. *et al.* A *de novo* paradigm for mental retardation. *Nat. Genet.* **42**, 1109–1112 (2010).
The first study to use exome sequencing of patient–parent trios to identify *de novo* mutations in a complex trait characterized by extreme genetic heterogeneity.
9. Yang, Y. *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* **369**, 1502–1511 (2013).
10. The Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **519**, 223–228 (2015).
First results of one of the largest sequencing projects to date, applying a statistical framework for the identification of new disease genes for developmental disorders.
11. Najmabadi, H. *et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478**, 57–63 (2011).
A first systematic study of the genetics underlying recessive intellectual disability using homozygosity mapping and targeted NGS to identify many new candidate ID genes.

12. Musante, L. & Ropers, H. H. Genetics of recessive cognitive disorders. *Trends Genet.* **30**, 32–39 (2014).
13. Moeschler, J. B., Shevell, M. & Committee on Genetics. Comprehensive evaluation of the child with intellectual disability or global developmental delays. *Pediatrics* **134**, e903–e918 (2014).
14. Lejeune, J., Turpin, R. & Gautier, M. Chromosomic diagnosis of mongolism. *Arch. Fr. Pediatr.* **16**, 962–963 (in French) (1959).
15. Lubs, H. A. A marker X chromosome. *Am. J. Hum. Genet.* **21**, 231–244 (1969).
16. Michelson, D. J. *et al.* Evidence report: genetic and metabolic testing on children with global developmental delay: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology* **77**, 1629–1635 (2011).
17. van Karnebeek, C. D., Jansweijer, M. C., Leenders, A. G., Offringa, M. & Hennekam, R. C. Diagnostic investigations in individuals with mental retardation: a systematic literature review of their usefulness. *Eur. J. Hum. Genet.* **13**, 6–25 (2005).
18. Pieretti, M. *et al.* Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell* **66**, 817–822 (1991).
19. Verkerk, A. J. *et al.* Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 (1991).
20. Coffee, B. *et al.* Incidence of fragile X syndrome by newborn screening for methylated *FMR1* DNA. *Am. J. Hum. Genet.* **85**, 503–514 (2009).
21. Leonard, H. & Wen, X. The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment. Retard. Dev. Disabil. Res. Rev.* **8**, 117–134 (2002).
22. Lubs, H. A., Stevenson, R. E. & Schwartz, C. E. Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* **90**, 579–590 (2012).
23. Vissers, L. E. & Stankiewicz, P. Microdeletion and microduplication syndromes. *Methods Mol. Biol.* **838**, 29–75 (2012).
24. Albertson, D. G. & Pinkel, D. Genomic microarrays in human genetic disease and cancer. *Hum. Mol. Genet.* **12**, R145–R152 (2003).
25. de Vries, B. B. *et al.* Diagnostic genome profiling in mental retardation. *Am. J. Hum. Genet.* **77**, 606–616 (2005).
26. Shaw-Smith, C. *et al.* Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J. Med. Genet.* **41**, 241–248 (2004).
27. Vissers, L. E. *et al.* Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am. J. Hum. Genet.* **73**, 1261–1270 (2003).
28. Wagenstaller, J. *et al.* Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am. J. Hum. Genet.* **81**, 768–779 (2007).
29. Cooper, G. M. *et al.* A copy number variation morbidity map of developmental delay. *Nat. Genet.* **43**, 838–846 (2011).
30. Koolen, D. A. *et al.* Mutations in the chromatin modifier gene *KANSL1* cause the 17q21.31 microdeletion syndrome. *Nat. Genet.* **44**, 639–641 (2012).
31. Lupski, J. R. & Stankiewicz, P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* **1**, e49 (2005).
32. Miller, D. T. *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* **86**, 749–764 (2010).
Important consortium statement on the use of genomic microarrays as first-tier diagnostic tests in individuals with intellectual disability and congenital abnormalities.
33. Schuurs-Hoeijmakers, J. H. *et al.* Homozygosity mapping in outbred families with mental retardation. *Eur. J. Hum. Genet.* **19**, 597–601 (2011).
34. Ng, S. B. *et al.* Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.* **42**, 30–35 (2010).
35. Hoischen, A. *et al. De novo* mutations of *SETBP1* cause Schinzel–Giedion syndrome. *Nat. Genet.* **42**, 483–485 (2010).
36. Ng, S. B. *et al.* Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
37. Hoischen, A. *et al. De novo* nonsense mutations in *ASXL1* cause Bohring–Opitz syndrome. *Nat. Genet.* **43**, 729–731 (2011).
38. Boycott, K. M., Vanstone, M. R., Bulman, D. E. & MacKenzie, A. E. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nat. Rev. Genet.* **14**, 681–691 (2013).
39. Keller, M. C. & Miller, G. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav. Brain Sci.* **29**, 385–404; discussion 405–452 (2006).
40. Hamdan, F. F. *et al.* Mutations in *SYNGAP1* in autosomal nonsyndromic mental retardation. *N. Engl. J. Med.* **360**, 599–605 (2009).
One of the first studies investigating the role of *de novo* mutations of candidate ID genes.
41. Hamdan, F. F. *et al. De novo* mutations in *FOXP1* in cases with intellectual disability, autism, and language impairment. *Am. J. Hum. Genet.* **87**, 671–678 (2010).
42. Veltman, J. A. & Brunner, H. G. *De novo* mutations in human genetic disease. *Nat. Rev. Genet.* **13**, 565–575 (2012).
43. Epi, K. C. *et al. De novo* mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
44. Fromer, M. *et al. De novo* mutations in schizophrenia implicate synaptic networks. *Nature* **506**, 179–184 (2014).
45. Girard, S. L. *et al.* Increased exonic *de novo* mutation rate in individuals with schizophrenia. *Nat. Genet.* **43**, 860–863 (2011).
46. Gulsuner, S. *et al.* Spatial and temporal mapping of *de novo* mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013).
47. Neale, B. M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
48. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485**, 246–250 (2012).
This study was one of the first of a series of large-scale trio-based sequencing projects identifying *de novo* mutations in neurodevelopmental disorders. The authors performed exome sequencing of more than 200 ASD patients and their parents.
49. Xu, B. *et al.* Exome sequencing supports a *de novo* mutational paradigm for schizophrenia. *Nat. Genet.* **43**, 864–868 (2011).
50. Hamdan, F. F. *et al. De novo* mutations in moderate or severe intellectual disability. *PLoS Genet.* **10**, e1004772 (2014).
51. Bernkopf, M. *et al.* Disruption of the methyltransferase-like 23 gene *METTL23* causes mild autosomal recessive intellectual disability. *Hum. Mol. Genet.* **23**, 4015–4023 (2014).
52. Iqbal, Z. *et al.* Homozygous *SLC6A17* mutations cause autosomal-recessive intellectual disability with progressive tremor, speech impairment, and behavioral problems. *Am. J. Hum. Genet.* **96**, 386–396 (2015).
53. Schuurs-Hoeijmakers, J. H. *et al.* 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 (2012).
54. MacArthur, D. G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469–476 (2014).
These authors proposed guidelines for summarizing confidence in variant pathogenicity in the era of NGS. The authors especially emphasize the use of statistical frameworks to assess genetic, informatic and functional evidence.
55. Gratten, J., Visscher, P. M., Mowry, B. J. & Wray, N. R. Interpreting the role of *de novo* protein-coding mutations in neuropsychiatric disease. *Nat. Genet.* **45**, 234–238 (2013).
56. Samocha, K. E. *et al.* A framework for the interpretation of *de novo* mutation in human disease. *Nat. Genet.* **46**, 944–950 (2014).
57. Coe, B. P. *et al.* Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* **46**, 1063–1071 (2014).
The largest study to date for copy number analysis in individuals with developmental delay.
58. Piton, A., Redin, C. & Mandel, J. L. XLID-causing mutations and associated genes challenged in light of data from large-scale human exome sequencing. *Am. J. Hum. Genet.* **93**, 368–383 (2013).
59. O'Roak, B. J. *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**, 1619–1622 (2012).
60. Stessman, H. A., Bernier, R. & Eichler, E. E. A genotype-first approach to defining the subtypes of a complex disease. *Cell* **156**, 872–877 (2014).
Important essay promoting a genome-first approach in which exome and genome sequencing studies identify candidate disease genes that are further studied in large case–control cohorts. Detailed clinical re-phenotyping is essential of all individuals with mutations in the same candidate disease gene to establish genotype–phenotype correlations.
61. Helsmoortel, C. *et al.* A SWI/SNF-related autism syndrome caused by *de novo* mutations in *ADNP*. *Nat. Genet.* **46**, 380–384 (2014).
62. Bernier, R. *et al.* Disruptive *CHD8* mutations define a subtype of autism early in development. *Cell* **158**, 263–276 (2014).
63. Zeidler, S., Hukema, R. K. & Willemsen, R. The quest for targeted therapy in fragile X syndrome. *Expert Opin. Ther. Targets* **19**, 1277–1281 (2015).
64. Davis, E. E., Frangakis, S. & Katsanis, N. Interpreting human genetic variation with *in vivo* zebrafish assays. *Biochim. Biophys. Acta* **1842**, 1960–1970 (2014).
65. Bellen, H. J., Tong, C. & Tsuda, H. 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. *Nat. Rev. Neurosci.* **11**, 514–522 (2010).
66. Oortveld, M. A. *et al.* Human intellectual disability genes form conserved functional modules in *Drosophila*. *PLoS Genet.* **9**, e1003911 (2013).
67. van der Voet, M., Nijhof, B., Oortveld, M. A. & Schenck, A. *Drosophila* models of early onset cognitive disorders and their clinical applications. *Neurosci. Biobehav. Rev.* **46**, 326–342 (2014).
68. Cundiff, P. E. & Anderson, S. A. Impact of induced pluripotent stem cells on the study of central nervous system disease. *Curr. Opin. Genet. Dev.* **21**, 354–361 (2011).
69. Wade, M. High-throughput silencing using the CRISPR–Cas9 system: a review of the benefits and challenges. *J. Biomol. Screen.* **20**, 1027–1039 (2015).
70. Krumm, N., O'Roak, B. J., Shendure, J. & Eichler, E. E. A *de novo* convergence of autism genetics and molecular neuroscience. *Trends Neurosci.* **37**, 95–105 (2014).
71. Snijders Blok, L. *et al.* Mutations in *DDX3X* are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. *Am. J. Hum. Genet.* **97**, 343–352 (2015).
72. Santen, G. W. *et al.* Mutations in SWI/SNF chromatin remodeling complex gene *ARID1B* cause Coffin–Siris syndrome. *Nat. Genet.* **44**, 379–380 (2012).
The recent identification of a novel ID gene for which mutations explain a relatively large fraction of ID patients. This study shows that frequently mutated ID genes can still be identified.
73. Oti, M., Huynen, M. A. & Brunner, H. G. Phenome connections. *Trends Genet.* **24**, 103–106 (2008).
74. van Bokhoven, H. Genetic and epigenetic networks in intellectual disabilities. *Annu. Rev. Genet.* **45**, 81–104 (2011).
75. Rothwell, P. E. *et al.* Autism-associated neuroligin-3 mutations commonly impair striatal circuits to boost repetitive behaviors. *Cell* **158**, 198–212 (2014).
76. Schubbert, S., Bollag, G. & Shannon, K. Deregulated Ras signaling in developmental disorders: new tricks for an old dog. *Curr. Opin. Genet. Dev.* **17**, 15–22 (2007).
77. Ba, W., van der Raadt, J. & Nadif Kasri, N. Rho GTPase signaling at the synapse: implications for intellectual disability. *Exp. Cell Res.* **319**, 2368–2374 (2013).
78. Loveday, C. *et al.* Mutations in the PP2A regulatory subunit B family genes *PPP2R5B*, *P2R5C* and *PPP2R5D* cause human overgrowth. *Hum. Mol. Genet.* **24**, 4775–4779 (2015).
79. Willemsen, M. H. *et al.* Update on Kleefstra syndrome. *Mol. Syndromol.* **2**, 202–212 (2012).
80. Forbes, S. A. *et al.* The catalogue of somatic mutations in cancer (COSMIC). *Curr. Protoc. Hum. Genet.* http://dx.doi.org/10.1002/0471142905.hg1011s57 (2008).

81. Hormozdiari, F., Penn, O., Borenstein, E. & Eichler, E. E. The discovery of integrated gene networks for autism and related disorders. *Genome Res.* **25**, 142–154 (2015).
82. Cristino, A. S. *et al.* Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. *Mol. Psychiatry* **19**, 294–301 (2014).
83. Hoischen, A., Krumm, N. & Eichler, E. E. Prioritization of neurodevelopmental disease genes by discovery of new mutations. *Nat. Neurosci.* **17**, 764–772 (2014).
84. Fombonne, E. Epidemiological trends in rates of autism. *Mol. Psychiatry* **7**, S4–S6 (2002).
85. Ronemus, M., Iossifov, I., Levy, D. & Wigler, M. The role of *de novo* mutations in the genetics of autism spectrum disorders. *Nat. Rev. Genet.* **15**, 133–141 (2014).
86. Jiang, Y. H. *et al.* Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am. J. Hum. Genet.* **93**, 249–263 (2013).
87. Liao, Y. *et al. SCN2A* mutation associated with neonatal epilepsy, late-onset episodic ataxia, myoclonus, and pain. *Neurology* **75**, 1454–1458 (2010).
88. McCarthy, S. E. *et al. De novo* mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol. Psychiatry* **19**, 652–658 (2014).
89. Nava, C. *et al. De novo* mutations in *HCN1* cause early infantile epileptic encephalopathy. *Nat. Genet.* **46**, 640–645 (2014).
90. Noebels, J. Pathway-driven discovery of epilepsy genes. *Nat. Neurosci.* **18**, 344–350 (2015).
91. Poduri, A., Evrony, G. D., Cai, X. & Walsh, C. A. Somatic mutation, genomic variation, and neurological disease. *Science* **341**, 1237758 (2013).
92. Acuna-Hidalgo, R. *et al.* Post-zygotic point mutations are an underrecognized source of *de novo* genomic variation. *Am. J. Hum. Genet.* **97**, 67–74 (2015).
93. Campbell, I. M. *et al.* Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am. J. Hum. Genet.* **95**, 173–182 (2014).
94. Riviere, J. B. *et al. De novo* germline and postzygotic mutations in *AKT3*, *PIK3R2* and *PIK3CA* cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* **44**, 934–940 (2012).
95. Jamuar, S. S. *et al.* Somatic mutations in cerebral cortical malformations. *N. Engl. J. Med.* **371**, 733–743 (2014).
96. Lindhurst, M. J. *et al.* A mosaic activating mutation in *AKT1* associated with the Proteus syndrome. *N. Engl. J. Med.* **365**, 611–619 (2011).
97. Cai, X. *et al.* Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep.* **8**, 1280–1289 (2014).
98. McConnell, M. J. *et al.* Mosaic copy number variation in human neurons. *Science* **342**, 632–637 (2013). **Important study for ID genetics in that it pioneers the detection of mosaic mutations in human neurons.**
99. Schuurs-Hoeijmakers, J. H. *et al.* Recurrent *de novo* mutations in *PACS1* cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am. J. Hum. Genet.* **91**, 1122–1127 (2012).
100. Mefford, H. C. *et al.* Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.* **359**, 1685–1699 (2008).
101. Sharp, A. J. *et al.* A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat. Genet.* **40**, 322–328 (2008).
102. Girirajan, S. *et al.* A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat. Genet.* **42**, 203–209 (2010).
103. Jacquemont, S. *et al.* A higher mutational burden in females supports a 'female protective model' in neurodevelopmental disorders. *Am. J. Hum. Genet.* **94**, 415–425 (2014).
104. Cooper, D. N. *et al.* Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum. Mutat.* **31**, 631–655 (2010).
105. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
106. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene–enhancer interactions. *Cell* **161**, 1012–1025 (2015). **A nice demonstration of the identification of *de novo* non-coding CNVs causing a Mendelian disorder. The authors show the extensive amount of functional evidence that is required to prove pathogenicity for non-coding variants.**
107. Weedon, M. N. *et al.* Recessive mutations in a distal *PTF1A* enhancer cause isolated pancreatic agenesis. *Nat. Genet.* **46**, 61–64 (2014).
108. Huang, L. *et al.* A noncoding, regulatory mutation implicates *HCFC1* in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **91**, 694–702 (2012).
109. Spielmann, M. & Mundlos, S. Structural variations, the regulatory landscape of the genome and their alteration in human disease. *Bioessays* **35**, 533–543 (2013).
110. Bredenoord, A. L., de Vries, M. C. & van Delden, J. J. Next-generation sequencing: does the next generation still have a right to an open future? *Nat. Rev. Genet.* **14**, 306 (2013).
111. Kong, A. *et al.* Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
112. Picker, J. D. & Walsh, C. A. New innovations: therapeutic opportunities for intellectual disabilities. *Ann. Neurol.* **74**, 382–390 (2013).
113. Miyake, N. *et al. KDM6A* point mutations cause Kabuki syndrome. *Hum. Mutat.* **34**, 108–110 (2013).
114. Bjornsson, H. T. *et al.* Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci. Transl. Med.* **6**, 256ra135 (2014). **Exciting study providing the first evidence that a deficiency in postnatal neurogenesis underlies the ID observed in Kabuki syndrome, and suggesting that this can be rescued by HDAC inhibitors.**
115. Berdasco, M. & Esteller, M. Genetic syndromes caused by mutations in epigenetic genes. *Hum. Genet.* **132**, 359–383 (2013).
116. Braat, S. & Kooy, R. F. The $GABA_A$ receptor as a therapeutic target for neurodevelopmental disorders. *Neuron* **86**, 1119–1130 (2015).
117. Na, E. S., Morris, M. J., Nelson, E. D. & Monteggia, L. M. $GABA_A$ receptor antagonism ameliorates behavioral and synaptic impairments associated with MeCP2 overexpression. *Neuropsychopharmacology* **39**, 1946–1954 (2014).
118. Leigh, M. J. *et al.* A randomized double-blind, placebo-controlled trial of minocycline in children and adolescents with fragile X syndrome. *J. Dev. Behav. Pediatr.* **34**, 147–155 (2013).
119. Verpelli, C., Galimberti, I., Gomez-Mancilla, B. & Sala, C. Molecular basis for prospective pharmacological treatment strategies in intellectual disability syndromes. *Dev. Neurobiol.* **74**, 197–206 (2014).
120. Knight, S. J. *et al.* Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* **354**, 1676–1681 (1999).
121. Bundey, S., Thake, A. & Todd, J. The recurrence risks for mild idiopathic mental retardation. *J. Med. Genet.* **26**, 260–266 (1989).
122. Lamont, M. A. & Dennis, N. R. Aetiology of mild mental retardation. *Arch. Dis. Child* **63**, 1032–1038 (1988).

Acknowledgements
We thank Hans-Hilger Ropers and Arjan de Brouwer for fruitful scientific discussions, and Rolph Pfundt for discussion and expertise on diagnostic relevant ID genes. This work was in part financially supported by grants from the Netherlands Organization for Scientific Research (916-14-043 to C.G. and 918-15-667 to J.A.V.) and the European Research Council (ERC; starting grant DENOVO 281964 to J.A.V.).

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (table)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF