Review

Making headway with genetic diagnostics of intellectual disabilities


Until recently, the cause of intellectual disability (ID) remained unexplained in at least 50% of affected individuals. Recent advances in genetic technologies led to great new opportunities to elucidate genetic defects implicated in ID. The introduction of genome-wide technologies that are able to detect small chromosomal copy number variations led to the identification of several microdeletion/duplication syndromes and to the subsequent identification of single causative genes. By the recent implementation of whole exome sequencing (WES) in research and diagnostics, with the potential to identify disease causing variants throughout the human exome at the base-pair level, a new revolution has started. Several studies showed that WES is effective in the identification of ID genes. Here we provide an historical overview of the advances in diagnostics of ID and illustrate the high diagnostic potential of current technologies by presenting the diagnostic survey that we performed in a series of 253 individuals with previously unexplained ID. This is the first study that systematically evaluated the diagnostic yield of the currently available and rapidly developing genetic diagnostic technologies lead to a significant increase in the number of patients that can be diagnosed.

M.H. Willemsen and T. Kleefstra
Department of Human Genetics,
Radboud University Medical Centre,
Nijmegen, The Netherlands

Key words: advances – diagnostic yield – genetic diagnostics – intellectual disability

Corresponding author: Tjitske Kleefstra, Department of Human Genetics, Radboud University Medical Centre, 6500 HB Nijmegen, The Netherlands. Tel.: +31-(0)243613946; fax: 0031-(0)243668753; e-mail: t.kleefstra@gen.umcn.nl [PO BOX 9101]

Received 29 April 2013, revised and accepted for publication 24 July 2013

Intellectual disability or mental retardation, what is in a name?

The nomenclature of intellectual disability (ID) has frequently been an issue of debate. The terminology has changed many times over the years. These name changes reflect the tendency in the attitude of the society to use more respectful and less offensive terms (1–6). ID and mental retardation (MR) are different descriptions of the same phenomenon. The term ‘mental retardation’ was introduced around 1960 (1, 2). During the last decade, the term ‘intellectual disability’ is increasingly being used and actually has generally replaced the term ‘mental retardation’ (3, 4). This transition in nomenclature was initiated by organizations that provide health care and social support to people with ID and increasingly gained ground in other areas, such as public authorities and more recently in (clinical) genetics as well. In 2006, the American Association on Mental Retardation (AAMR) changed her name to the American Association on Intellectual and Developmental Disabilities (AAIDD) (4). Furthermore, in October 2010 President Obama signed ‘Rosa’s law’, mandating that the term ‘mental retardation’ should be replaced by ‘intellectual disability’ in the federal statutes (http://blogs.suntimes.com/sweet/2010/10/obama_signs_rosas_law_mental_r.html).

Opponents of the name change state that the new name is less accurate, because it does not indicate that it is a developmental problem, and that it causes confusion, because it is not associated with a change in the definition (5). Most international classification systems, such as the Diagnostic and Statistical Manual of Mental Disorders, Fourth revision (DSM-IV) of the American Psychiatric Association and the International Classification of Diseases, Tenth Revision (ICD-10) of the World Health Organization (WHO) are based on the definition by the AAIDD. This definition comprises three criteria: (i) a significant limitation in intellectual functioning, (ii) a significant limitation in adaptive
behavior, like social and practical skills, and (iii) an origin before the age of 18 years (4).

**Intellectual disability, an issue of major personal and social impact**

On the basis of an assumed normal distribution of IQ values in the general population, 2–3% of the individuals have an IQ level ≤70. However, the distribution of intelligence is not normally distributed and in fact represented by two distribution curves, a normal distribution curve, with a mean IQ of 100, and a second curve with a mean of about 35, representing individuals with ID because of pathophysiological defects. Mild forms of ID, especially those with IQ levels around 70, are thought to represent the lower end of the normal IQ distribution and to be more multifactorial (7, 8).

In previous studies, prevalence estimations of ID range from 0.2% to 8%. This broad range is caused by differences in study design, study population and definition of ID (8–11). The reported prevalence of severe ID (IQ < 50) is lower than the prevalence of mild ID (IQ > 50) and is less variable. Most studies report a prevalence of severe ID in the range of 0.3–0.5% (8–12).

ID is one of the main reasons for referral to a clinical geneticist and has a major impact on affected individuals, their families and society. Most individuals with ID have a lifelong need of care and they have significantly more physical and mental co-morbidity (13–15). This makes ID one of the most costly health issues.

**A genetic diagnosis makes the difference**

Knowing the cause of ID is of great importance. It may provide insight in co-morbidity, associated behavior problems, prognosis and lifespan, and recurrence risk, and thereby gives answers to important questions of involved families and health care providers. Identification of a genetic cause precludes further unnecessary and often incriminating testing, and fruitless interventions. Instead, specific anticipation on associated health and behavioral problems is facilitated, since the extreme heterogeneity in etiology and clinical presentation goes along with diverse prognoses and variable needs. Besides, the emotional impact on affected individuals and/or their families of just getting an explanation is often enormous. In addition, a diagnosis enables adequate counseling of recurrence risk and prenatal testing. Moreover, the identification of genetic defects associated with ID provides insight in the underlying pathological mechanisms, which is the first crucial step that opens the way toward the future development of treatment strategies (16, 17).

**Chromosomal aberrations**

In up to 25% of the individuals with ID, causal chromosomal aberrations can be identified, including both microscopically visible and submicroscopic aberrations. Microscopically visible chromosomal aberrations can be detected in 12.5–14.5% (8% + 4.5–6.5%) of which trisomy 21 is the most frequent, accounting for about 8% (8, 19–24, 26, 27).

The introduction of targeted chromosomal analysis by Fluorescent In Situ Hybridization (FISH) and multiplex ligation dependent probe amplification (MLPA) enabled the detection of microdeletions/duplications in patients with clinically recognizable syndromes, such as Prader-Willi and Angelman Syndrome and Williams syndrome, and (subtelomeric) copy number variations (CNVs) in other individuals by the use of (sub) telomere probes. The application of these techniques led to the identification of a molecular diagnosis in an additional 6–10% (Table S1) (8, 23, 27–29).

The subsequent introduction of genome-wide array analysis over the last decade allowed detection of (sub)microscopic CNVs with an increasing improvement in resolution up to the size of a single exon and led to the identification of numerous novel microdeletion and microduplication syndromes, which previously escaped detection by routine (molecular) cytogenetic
techniques (18, 30–34). The phenotypic characteristics of some microdeletion syndromes were subsequently found to be caused by haploinsufficiency of single genes, such as \textit{EHMT1} in Kleefstra syndrome (35), \textit{TCF4} in Pitt-Hopkins syndrome (36) and \textit{KANSL1} in Koolen-deVries syndrome (37, 38). Moreover, this application also enabled gene identification of numerous other neurodevelopmental disorders (39, 40). After the introduction of genome-wide array technologies, several studies on the diagnostic yield have been reported, reviewed by Hochstenbach et al. (31), Sagoo et al. (41) and Miller et al. (30). Together these studies showed an average diagnostic yield of 15–20%, including microscopically visible chromosomal aberrations, but excluding trisomy 21.

Monogenic causes

Dominant ID

Severe AD ID is almost always of \textit{de novo} origin. It is assumed that \textit{de novo} copy number variations (CNVs) and point mutations may explain the majority of severe ID (8, 42, 43). In the past 3 years, several studies showed that whole exome sequencing (WES) is an effective approach for the identification of \textit{de novo} mutations in clinically recognizable syndromes (44–46). Moreover, WES is also effectively applied in the identification of \textit{de novo} mutations through a family-based approach (child–parents trio’s) in individuals with sporadic ID without a clinically recognizable syndrome (42, 43, 47, 48). In a series of 100 persons with severe ID (IQ < 50), we identified in 10 individuals \textit{de novo} mutations in known AD ID genes, in three mutations in known X-linked ID genes, in three \textit{de novo} mutations in novel identified ID genes and in 19 individuals mutations in potentially novel ID genes, leading to a (potential) diagnosis in 35% in total (43). Rauch et al. (47) concluded from their study that about 45–55% of the cases of severe ID could be explained by \textit{de novo} mutations [including both (possible) damaging mutations in known and novel (candidate) ID genes]. Furthermore, identification of \textit{de novo} mutations in novel ID genes in individuals with overlapping ID phenotypes has led us to define novel clinical syndromes (49–51).

Studies in cohorts of individuals with autism showed that \textit{de novo} mutations account for a substantial part of autism as well. Both clinical and molecular findings show a remarkable overlap between ‘autism cohorts’ and ‘ID cohorts’ suggesting common biological (etiological) pathways (52, 53).

Recessive ID

It is estimated that the total number of genes involved in AR ID may well run into the thousands (8, 54). Only a very small minority of these genes has been elucidated by now due to the lack of suitable families to be studied in Western countries (8, 54, 55). Homozygosity mapping in large consanguineous families was the approach for localizing AR ID genes and numerous loci for AR ID were reported. By combining this approach with targeted sequencing of genes in candidate regions, several AR ID genes were elucidated (8, 55–58).

Exome sequencing has now also been shown successful in the identification of AR ID genes. Several publications reported the identification of the responsible gene in consanguineous ID families (59–63). In addition, exome sequencing in syndromic forms of AR ID in multiple non-consanguineous families led to the elucidation of the underlying gene defects in AR ID syndromes, such as Dubowitz syndrome (OMIM 223370) and CHIME syndrome (OMIM 280000) (64, 65). Najmabadi et al. (54) used a combination of homozygosity mapping and NGS after targeted enrichment of exons from homozygous linkage intervals in a large series of 136 consanguineous ID
## Willemsen and Kleefstra

### Table 1. Yield of previous systematic cohort studies on the etiology of intellectual disability (ID) before the genome-wide array analysis era

<table>
<thead>
<tr>
<th>Study</th>
<th>Patient cohort</th>
<th>Cytogenetic diagnoses</th>
<th>Conventional karyotyping</th>
<th>TRISOMY 21</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>471, mainly male patients</td>
<td>21.2%</td>
<td>21%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(age range 3–75 years, IQ &lt; 50 in &gt;80%)</td>
<td>11%</td>
<td>~11%</td>
<td>2.9% (in small subset subtelomere anomalies)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10,997 patients, 59% males</td>
<td>11%</td>
<td>7.8%</td>
<td>7.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>57% &gt;20 years, IQ &lt; 50</td>
<td>'overrepresented'</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34.2%</td>
<td>21.2%</td>
<td>11%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32%</td>
<td>21%</td>
<td>~11%</td>
<td>2.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34.2%</td>
<td>32%</td>
<td>21%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.6%</td>
<td>21%</td>
<td>17%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>47.8% (b)</td>
<td>38% (f)</td>
<td>30.6% (h)</td>
<td>40.8%</td>
<td></td>
</tr>
</tbody>
</table>

**Study** | **Patient cohort** | **Cytogenetic diagnoses** | **Conventional karyotyping** | **TRISOMY 21** | **Other**
---|---|---|---|---|---
| Van Buggenhout et al. 2001 (22) | 471, mainly male patients | 21.2% | 21% | 3% |
| Stevenson et al. 2003 (20) | 10,997 patients, 59% males | 11% | ~11% | 2.9% |
| Van Karnebeek et al. 2005 (19) | 281 patients, 52% male patients, all children | 11% | 7.8% | 7.8% |
| Rauch et al. 2006 (21) | 670, mostly with an age < 6 years, male: female ratio not mentioned | 30.7% | 17.5% | 9.2% |

**FISH/MLPA** | **Molecular cytogenetic** | **Monogenic diagnoses** | **Fragile X** | **Other** | **(Subtelomere) FISH/MLPA** |
---|---|---|---|---|---
| – | – | – | – | – |
| – | – | – | – | – |
| 13% | 8% (d) | 2.8% | 4.8% | 2% |
| 3.2% | 1.7% | 0.4% | 1.2% | 13.9% |
| 9.8% (a) | 5.3% (d) | 2.4% | 3.6% | 18% |
| (a) | Not mentioned (e) | 1.8% (g) | Not mentioned (j) | – |
| (a) | 2% | 1.7% | 1.1% | 2% |
| 34.2% | 21% | 29.5 | 39.5 | 40.8% |
| 14.6% | 17% | 1.1% | 4% | 13.9% |
| 47.8% (b) | 38% (f) | 30.6% (h) | 40.8% | 40.8% |

- **FISH**, Fluorescent In Situ Hybridization; **MLPA**, multiplex ligation dependent probe amplification.
- (a), metabolic diagnoses (3.4%) and clinical diagnoses were included in the group of monogenic diagnoses; (b), excluding 1.7% diagnosed with a CNS malformation; (c), the percentages of the subgroups numerical, structural and subtelomere anomalies are not further quantified, together they amount to a total of 2.9%; (d), it is not indicated whether all cases were molecularly confirmed. In addition microdeletions associated with Prader-Willi, Angelman, Williams an velocardiofacial syndrome were included in this category; (e), likely these were included in the category monogenic disorders; (f), diagnostic categories ‘multifactorial’ and ‘culturofamilial’ were included in this category or in the category monogenic disorders. The category ‘Mendelian disorders’ was not taken into account here, because the diagnoses in this category were not molecularly confirmed.

Families. They reported in 54% potentially disease-causing variants in known and novel candidate ID genes. Reports of the identification of AR ID genes in non-syndromic outbred families by exome sequencing approaches are more rare (66, 67). Analysis of WES data for AR inheritance in single cases of ID, suggests that AR ID is rare in individuals with isolated ID and non-consanguineous parents (43, 47).

### X-linked (XL) ID

Over 100 genes linked to ID so far are located on the X-chromosome and are thought to explain most of the excess of ID in males. Estimations of the proportion of ID in males that is XL range from 5% to 12% (8, 10, 68).

Fragile X syndrome is after trisomy 21 by far the most common cause of ID with a prevalence of approximately 0.5–3% in different populations of individuals with ID (19–22).

Historically, gene discovery mainly focused on the X-chromosome (8, 68). Large international collaborations, such as EURO-MRX (http://www.euromrx.com), have played a major role. Lubs et al. reviewed the success of the main strategies before the NGS era, of which positional cloning based on a chromosome rearrangement, and linkage studies in combination with candidate gene screening were most successful (8, 68, 69).

The majority of previously published genome-wide array studies in unselected cohorts of patients with ID, CA and/or dysmorphic features reported a detection rate of X-chromosome CNVs (X-CNVs) of 0–5%, but this varied from 0% to 40% because of differences in patient selection and array platforms. Studies using X-chromosome specific array-CGH reported a higher detection rate of X-CNVs which is also partly explained in patient selection and array platforms. Studies using X-chromosome specific array-CGH reported a higher detection rate of X-CNVs which is also partly explained by the selection of a cohort of predominantly male patients with the suspicion XL inheritance (reviewed in Ref. (70)).

Gene identification through study of X-CNVs only led to the identification of few ID genes, as HUWE1 and ZNF674. The most frequent pathogenic X-CNVs are Xq28 duplications encompassing MECP2 (68, 71).
Sequencing of all genes on the X-chromosome led to the identification of 9 novel XL ID genes (72). Targeted sequencing of 86 known XL genes resulted in the detection of potentially disease-causing sequence variants in 42% of X-linked families (excluding Fragile X syndrome) (73). Kalscheuer et al. estimated that up to 70% of X-linked ID can be explained by sequencing of all exonic sequences on the X-chromosome, based on studies in >250 families with XL-ID (Kalscheuer et al., presented at ESHG meeting, Amsterdam 2011).

Yield of currently available genetic diagnostic repertoire in a diagnostic setting: a cohort study

The magnitude of the diagnostic impact of the total arsenal of recent advances in genetic diagnostic technologies has not yet been systematically assessed in etiologic surveys among large series of ID patients. We performed a systematic genetic diagnostic survey among a series of 253 individuals with unexplained ID, using the currently available clinical genetic diagnostic repertoire, thereby showing the diagnostic potential.

Methods

We selected 253 individuals from 234 families with unexplained ID originating from a cohort of 2069 individuals from Dutch service providers. Selection criteria and a flowchart of patient selection and recruitment are summarized in Table S2 and Figure S1. Detailed characteristics of the selected cohort are presented in Table S3. The parents or legal representatives gave their written informed consent. The study was approved by the local ethical committee (NL13636.091.07).

The study was divided into two phases such as (i) Diagnostic phase and (ii) Diagnostic-related research phase. During the first phase, we performed clinical investigations and used genetic diagnostic tests that were already routinely used. Specific DNA diagnostic tests were requested based on the clinical suspicion of a recognizable syndrome in a subset of the patients. Genome-wide microarray analysis and a metabolic screen were performed as standard. In all patients, except for those with microcephaly, DNA analysis of the FMR1 gene was performed. For further details see Methods in Appendix S1. Segregation of genetic/chromosomal variants was tested in the parents and/or other family members.

About 30% (58/191) of the undiagnosed individuals in the first study phase, were further analyzed by various NGS approaches in the second phase.

Forty-two isolated patients were considered for the involvement of de novo mutations with a dominant effect. These patients were potential candidates for family-based WES. Patients with an IQ < 50, normal karyotype, negative family history and availability of both parental DNA samples were selected for this approach, as previously described (42, 43). The analysis also allows the identification of recessive mutations, although these patients were selected against AR inheritance.

Sixteen index patients with familial ID (with/without homozygous regions ≥3 Mb) were included in further NGS studies including WES or X-exome sequencing to search for compound heterozygous, homozygous and/or X-linked variants. In three AR families, homozygosity mapping (74) was performed initially, followed by Sanger sequencing of candidate genes in the overlapping homozygous regions. Index patients from families with clear XL inheritance were included in X-exome sequencing studies in collaboration with the Max Planck Institute for Molecular Genetics in Berlin.

Validation and segregation analysis of putative causative variants were done by Sanger sequencing.

Results

In the first study phase, a diagnosis was established in 43 (18.4%) index individuals, including chromosomal aberrations (12.4%), monogenic defects (4.7%) and metabolic disorders (1.3%) (Table 2, Fig. 2, Tables S4 and S5).

In the second study phase a pertinent or plausible diagnosis was established in 25 additional cases, including in 24 of 58 (41.4%) individuals that received NGS analyses [19 individuals with sporadic ID (trio analysis) and 5 individuals with familial ID).

In 42 individuals trio-based exome sequencing was applied. The majority were previously described in the study of de Ligt et al. Nineteen (45.2%) of them received a (likely) molecular diagnosis (Table S6). Fourteen individuals have a pathogenic de novo mutation in a known or novel recurrent ID gene (08-0006, 09-0003, 09-0023, 09-0034, 09-0045, 09-0072, 09-0079, 10-0022, 10-0024, 10-0040, 10-0064, 10-0103, 10-0108, 11-0034). The remaining five have a de novo mutation in a likely novel ID gene, but additional cases have not yet been reported.

In 16 index individuals with familial ID, NGS was performed, either whole X chromosome sequencing, or WES. Five of them received a (likely) molecular diagnosis. In two families (10-0056 and 10-0076) mutations in known ID genes, segregating with the phenotype were detected. In one family (11-0006) heterozygous mutations in B3GALNT2, segregating with the phenotype in affected family members, was detected (Table S7). B3GALNT2 was recently identified as a novel causative gene for Walker–Warburg syndrome and dystroglycanopathies (75, 76). The phenotype in our family falls not within the clinical spectrum of dystroglycanopathies, therefore a causal relation is difficult to proof. In two families segregating mutations in plausible novel ID genes were identified. Additional functional studies (unpublished data) strongly supported involvement in ID phenotypes.

In addition to the two cases identified by WES, one additional patient with a mutation in the novel ID gene CTNNB1 (additional diagnosis number 25) was found by Sanger sequencing in a large confirmation series (43).
Table 2. General overview of abnormal findings in the diagnostic phase

<table>
<thead>
<tr>
<th>Abnormal finding</th>
<th>(Likely) pathogenic (a)</th>
<th>Unknown (a)</th>
<th>(Likely) non-pathogenic (a)</th>
<th>Total N (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosomal, CNV</td>
<td>29 (12.4)</td>
<td>16 (6.8)</td>
<td>11 (4.7)</td>
<td>56 (23.9)</td>
</tr>
<tr>
<td>Monogenic defect</td>
<td>11 (4.7)</td>
<td>0 (0)</td>
<td>3 (1.3)</td>
<td>14 (6.0)</td>
</tr>
<tr>
<td>Homozygous region (&gt;3 Mb)</td>
<td>0 (0) (b)</td>
<td>21 (9.0)</td>
<td>0 (0)</td>
<td>21 (9.0)</td>
</tr>
<tr>
<td>Metabolic abnormality</td>
<td>3 (1.3) (c)</td>
<td>18 (7.7)</td>
<td>66 (28.2)</td>
<td>87 (37.2)</td>
</tr>
<tr>
<td>Total N (%)</td>
<td>43 (18.4)</td>
<td>55 (23.5)</td>
<td>80 (34.2)</td>
<td>178 (76.1)</td>
</tr>
</tbody>
</table>

CNV, copy number variation.

*(a), number (%) of 234 index participants; (b), in one patient (09-0038) a ‘pathogenic’ homozygous region was found, which was explained by a mosaic trisomy. This patient was included in the chromosomal CNV group; (c), all mitochondrial disorders, not otherwise specified, because further investigations (in skin biopsy) were not possible.

Fig. 2. Clinical photographs of six individuals who were diagnosed in this survey. (a) Individual with 17q21.3 microdeletion (Koolen-deVries) syndrome (27 years). (b) Individual with 1p21.3 microdeletion (34 years). (c) Individual with 22q13.32q13.33 deletion (48 years). (d) Individual with Kleefstra syndrome (EHMT1 mutation) (41 years). (e) Individual with Pitt-Hopkins syndrome (TCF4 mutation) (41 years). (f) Individual with GATAD2B mutation (34 years).

Combining both study phases, the total yield of (likely) molecular diagnoses in our cohort of individuals with so far unexplained ID, sums up to 29.1% (68 of 234). Only 58 of the 191 patients (30.4%) who remained undiagnosed in the first phase were selected for NGS studies in the second phase. Taking this into account, the total diagnostic potential may be well estimated by the sum of the yield in both phases of the study, which is 59.8% (18.4 + 41.4%) (Fig. 3).

Discussion

Diagnostic potential

Hitherto, the yield of the recent impressive progress in genetic diagnostics has not been systematically explored in large series of individuals with ID. The present review aimed to assess the diagnostic potential of the currently available genetic diagnostic repertoire by including a clinical diagnostic survey in a cohort of individuals, who had not yet received modern clinical genetic diagnostic assessment. More than 18% received a (confirmed) diagnosis in the first study phase. Chromosomal aberrations accounted for the major part of diagnoses (12.4%), which is in line with the results of previous studies (16, 30, 41, 77).

The yield of metabolic screening in this study was very low (1.3%). This may be partly explained by the composition of the patient cohort, including mainly adults. Our findings are in line with previous studies. Therefore, selective and targeted metabolic testing is recommended (19, 78–80).

Despite some limitations, the impact of the application of present-day advanced genetic diagnostic technologies on the elucidation of genetic causes of unexplained ID may be roughly estimated by extrapolating the diagnostic yield of the present studies. We made some assumptions to calculate this estimation. First, we assumed that at least 50% of the cases of ID remained unexplained before the genome-wide array analysis era. Second, we considered it likely that this study cohort is composed of individuals from this group of 50% unexplained cases of ID. This assumption is supported by the fact that none of the participants had received genome-wide array analysis. Extrapolation of the yield in the diagnostic phase of this study (18.4%) to the total population of people with ID would then result in a 9.2% (18.4 × 0.5) increase in the total number of diagnosed patients to approximately 59% (9.2% + 50%). Extrapolation of the yield of additional NGS studies in the second study phase would lead to an additional 17% (41.4 × 0.41) increase in diagnoses to more than 75%. As these calculations are based on assumptions, several other remarks should be made. First, our study cohort (Table S3) may not be representative for the total population of people with ID. Overrepresentation of patients with an IQ < 50 might have led to an increased yield, although we did not observe a significant difference in the distribution of the level of ID between the group of diagnosed and undiagnosed patients (data not shown). Owing to our selection criteria individuals with additional findings to ID, such as facial dysmorphism, are overrepresented in our cohort. We only observed a significant difference in the presence of facial dysmorphism between the groups with and without diagnosis (significant more facial dysmorphism in patients with a diagnosis).
In the second study, phase only a subset of the patients was selected for NGS studies. Application of NGS in the total group of patients that remained undiagnosed in the first study phase, might have resulted in a lower diagnostic rate, because we have now excluded individuals with an IQ between 50 and 70 from family-based WES.

**Importance of widespread data sharing**

The application of WES in an increasing number of individuals will shed more light on the clinical significance of hitherto unknown variants. Some currently unknown variants may turn out to be either pathogenic or non-pathogenic. Recurrent findings in additional patient series and/or in vivo/in vitro functional studies may further support pathogenicity of mutations. Therefore, the diagnostic rate of exome sequencing will likely further increase. Moreover, additional novel ID syndromes will be defined as more recurrent mutations will be identified in patients with overlapping phenotypes. This highlights the importance of storage of molecular data in combination with detailed anonymized phenotypic data in central databases, in analogy to existing databases for the interpretation of CNVs, such as DECIPHER, ECARUCA, and DGV (81–85). These data may be integrated in the UCSC genome browser (http://genome.ucsc.edu).

**Detection of CNVs and mosaic mutations by NGS**

At present, in most clinical diagnostic centers array analysis is used as the first tier diagnostic test in individuals without a clinical recognizable ID syndrome. Ongoing progress in NGS technologies and bioinformatic analyses provides opportunities for the identification of CNVs in NGS data (both exome and whole genome) (86). Although sensitivity and consistency need to be thoroughly explored in a diagnostic setting, in the coming years, NGS will likely gradually replace genome-wide array analysis as the first tier diagnostic test. Major advantage is that a complete range of genetic aberrations, from point mutations to CNVs and aneuploidies can be detected by one single diagnostic test, which will eventually be more cost-effective and time-saving.

Somatic mosaicism for CNVs and intragenic mutations that lead to severe ID might go unnoticed due to inability of commonly used diagnostic tests to detect low level mosaicism, but may be more frequent than is generally assumed (87). Recently, several studies reported the detection of (low level) mosaic point mutations by NGS techniques in various disorders (88, 89). However, the sensitivity of the detection of (low level) point mutations and the frequency with which these occur in ID syndromes need to be further determined.

**Reverse phenotyping**

With the introduction of WES extensive phenotyping will become less important in guiding diagnostics, and the success of identification of clinical recognizable ID syndromes will no longer be highly dependent on clinical expertise on syndrome recognition (81). This will result in a shift from a ‘phenotype first’ to a ‘genotype first’ approach (reverse phenotyping). In the near future, mapping of the genomic profile may be the first step in the diagnostic process. Subsequent clinical investigations will be essential in the interpretation of detected genomic variants and to correlate these to the clinical presentation. Moreover, in-depth clinical investigations, including longitudinal investigations during follow-up will stay for management and insight in prognosis of individuals.
with a specific genetic defect. The results of such studies will be mandatory for future trials in which the phenotypic effects of drugs will be assessed.

As genetic causes of ID are very heterogeneous, each single cause is mostly very rare. Consequently, detailed clinical information is scarce. Establishment of outpatient clinics for diagnostics and follow-up in a multidisciplinary setting, bundling the expertise of the clinical geneticist and other medical specialties, enables further increase of knowledge and experience and facilitates research initiatives.

Major advantages of reverse phenotyping may be prevention of incriminating and invasive diagnostic tests and wrong/delayed diagnoses. Another consequence of the ‘genotype first’ approach will be a broadening of the clinical spectrum of known ID syndromes, because genetic defects may be detected in patients with less specific phenotypes.

**Whole genome sequencing**

Mutations outside the coding regions that cause ID remain to be defined. It is expected that the application of WES techniques in research and diagnostics of ID will be rapidly followed-up by whole genome sequencing (WGS) (81). This will enable the detection of disease causing variants in non-coding regulatory sequences, such as intronic variants, as well. These variants may lead to altered expression of genes and elucidate novel epigenetic ID causing mechanisms, although interpretation of the data will be in first instance complex and time-consuming. Epigenetic mechanisms have been shown to be involved in several ID syndromes (90).

**Therapeutic strategies**

It is supposed that the large number of ID genes converge onto a limited number of common underlying ID pathways, which may be the target of future therapeutic interventions. Of course, it appears unreasonable to expect that in patients with severe ID normal intellect can be restored, but targeted interventions in ID-related pathways might indeed improve cognitive, motor and social functioning and reduce degenerative aspects (17).

Recent studies have shown that there are indeed opportunities. Studies in animal models of Fragile X syndrome showed that metabotropic glutamate receptors 5 (mGluR5) antagonist and gamma amino-butyric acid (GABA) agonist can rescue various aspects of the phenotype by targeting molecular pathways disrupted in Fragile X syndrome (91, 92). Results of several clinical trials in humans have already been published and more studies are ongoing at the moment (reviewed in Ref. (91)). These studies showed promising results for pharmaco-therapeutical interventions that may not only improve behavioral aspects, but cognitive aspects as well (91).

For several ID genes, such as *FMR1*, *MECP2*, and *EHMT1*, it has been shown that the mutations give also rise to acute effects that can be rescued at adult stage, indicating that learning and memory are dynamic processes that might be rescued at least to some extent (91–94).

**Supporting Information**

The following Supporting information is available for this article:

- Fig S1. Flowchart of patient selection and recruitment.
- Appendix S1. Methods.
- Table S1. Historical hallmarks in genetic diagnostic technologies of ID.
- Table S2. Selection criteria cohort ID individuals.
- Table S3. Characteristics cohort ID individuals.
- Table S4. (Likely) pathogenic chromosomal copy number variations detected in the total cohort (N = 29).
- Table S5. Monogenic defects identified by specific DNA diagnostic tests in diagnostic phase of the study.
- Table S6. (Likely) pathogenic *de novo* mutations identified by WES in a total of 42 ID individuals.
- Table S7. Mutations in known ID genes identified by WES in familial ID.

Additional Supporting information may be found in the online version of this article.

**Acknowledgements**

We would like to thank the patients and their families. In addition, we thank professor H. van Bokhoven, professor B. C. J. Hamel (emeritus) and professor H. M. J. van Schrojenstein-Lantman de Valk for supervising. This work was supported by grants from the Consortium ‘Stronger on your own feet’ to T. K. and M. H. W., The Netherlands Organization for Health Research and Development (ZonMW grant 907-00-365 to T. K.) and the European Union under the 7th framework program (Gendecodys HEALTH-F4-2010-241995 to T. K.)

**Conflict of interest**

Nothing to declare.

**References**

Making headway with genetic diagnostics of intellectual disabilities


