Molecular Psychiatry (2017) 22, 615–624

www.nature.com/mp

ORIGINAL ARTICLE

# Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield

S Anazi[1,15], S Maddirevula[1,15], E Faqeih[2,15], H Alsedairy[1], F Alzahrani[1], HE Shamseldin[1], N Patel[1], M Hashem[1], N Ibrahim[1], F Abdulwahab[1], N Ewida[1], HS Alsaif[1], H Al sharif[1], W Alamoudi[1], A Kentab[3], FA Bashiri[3], M Alnaser[3], AH AlWadei[4], M Alfadhel[5], W Eyaid[5], A Hashem[6], A Al Asmari[2], MM Saleh[2], A AlSaman[4], KA Alhasan[3], M Alsughayir[7], M Al Shammari[3], A Mahmoud[4], ZN Al-Hassnan[1], M Al-Husain[3], R Osama Khalil[8,9], N Abd El.Meguid[9], A Masri[12], R Ali[13], T Ben-Omran[13], P El.Fishway[10], A Hashish[9], A Ercan Sencicek[10], M State[8], AM Alazami[1], MA Salih[3], N Altassan[1], ST Arold[11], M Abouelhoda[1], SM Wakil[1], D Monies[1], R Shaheen[1] and FS Alkuraya[1,14]

Intellectual disability (ID) is a measurable phenotypic consequence of genetic and environmental factors. In this study, we prospectively assessed the diagnostic yield of genomic tools (molecular karyotyping, multi-gene panel and exome sequencing) in a cohort of 337 ID subjects as a first-tier test and compared it with a standard clinical evaluation performed in parallel. Standard clinical evaluation suggested a diagnosis in 16% of cases (54/337) but only 70% of these (38/54) were subsequently confirmed. On the other hand, the genomic approach revealed a likely diagnosis in 58% ($n = 196$). These included copy number variants in 14% ($n = 54$, 15% are novel), and point mutations revealed by multi-gene panel and exome sequencing in the remaining 43% (1% were found to have Fragile-X). The identified point mutations were mostly recessive ($n = 117$, 81%), consistent with the high consanguinity of the study cohort, but also X-linked ($n = 8$, 6%) and *de novo* dominant ($n = 19$, 13%). When applied directly on all cases with negative molecular karyotyping, the diagnostic yield of exome sequencing was 60% (77/129). Exome sequencing also identified likely pathogenic variants in three novel candidate genes (*DENND5A*, *NEMF* and *DNHD1*) each of which harbored independent homozygous mutations in patients with overlapping phenotypes. In addition, exome sequencing revealed *de novo* and recessive variants in 32 genes (*MAMDC2, TUBAL3, CPNE6, KLHL24, USP2, PIP5K1A, UBE4A, TP53TG5, ATOH1, C16ORF90, SLC39A14, TRERF1, RGL1, CDH11, SYDE2, HIRA, FEZF2, PROCA1, PIANP, PLK2, QRFPR, AP3B2, NUDT2, UFC1, BTN3A2, TADA1, ARFGEF3, FAM160B1, ZMYM5*, *SLC45A1*, *ARHGAP33* and *CAPS2*), which we highlight as potential candidates on the basis of several lines of evidence, and one of these genes (*SLC39A14*) was biallelically inactivated in a potentially treatable form of hypermanganesemia and neurodegeneration. Finally, likely causal variants in previously published candidate genes were identified (*ASTN1*, *HELZ*, *THOC6*, *WDR45B, ADRA2B* and *CLIP1*), thus supporting their involvement in ID pathogenesis. Our results expand the morbid genome of ID and support the adoption of genomics as a first-tier test for individuals with ID.

*Molecular Psychiatry* (2017) **22,** 615–624; doi:10.1038/mp.2016.113; published online 19 July 2016

## INTRODUCTION

Intellectual disability (ID) is a common morbidity affecting at least 1% of the population, a fraction that represents the extreme end of the distribution curve of cognitive capacity in humans.[1] The natural history of ID is usually a stable course of impaired cognition presenting in early childhood as delayed acquisition of speech and other cognitive domains, and persisting into adulthood with variable degrees of limited mental function. Alternatively, ID may follow a progressive course characteristic of neurodegenerative diseases of childhood (ID only applies to the developing brain by definition[2]) where loss of cognitive skills follows a period of normal development.

The exact contribution of genetics to ID is unknown. Previous estimates relied on the recognition of identifiable genetic syndromes or positive family history even though the absence of these criteria is still compatible with a genetic etiology of ID. The extreme genetic heterogeneity of ID was a major impediment to the establishment of molecular diagnosis in the absence of a recognizable clinical syndrome or positional mapping data that guide the search for the likely candidate gene. The development

[1]Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; [2]Department of Pediatric Subspecialties, Children's Hospital, King Fahad Medical City, Riyadh, Saudi Arabia; [3]Department of Pediatrics, College of Medicine & King Khalid University Hospital, King Saud University, Riyadh, Saudi Arabia; [4]Pediatric Neurology Department, National Neuroscience Institute, King Fahad Medical City, Riyadh, Saudi Arabia; [5]Department of Pediatrics, King Saud bin Abdulaziz University for Health Science, King Abdulaziz Medical City, Riyadh, Saudi Arabia; [6]Department of Pediatrics, Prince Sultan Military Medical City, Riyadh, Saudi Arabia; [7]Department of Psychiatry, College of Medicine, King Saud University, Riyadh, Saudi Arabia; [8]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, USA; [9]National Research Center, Cairo, Egypt; [10]Department of Neurosurgery, Program on Neurogenetics, Yale University School of Medicine, New Haven, CT, USA; [11]King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Division of Biological and Environmental Sciences and Engineering (BESE), Thuwal, Saudi Arabia; [12]Department of Pediatrics, Faculty of Medicine, The University of Jordan, Amman, Jordan; [13]Clinical & Metabolic Genetics, Pediatrics, Hamad Medical Corporation, Doha, Qatar and [14]Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia. Correspondence: Dr R Shaheen or Professor FS Alkuraya, Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Saudi Arabia.
E-mail: rshaheen@kfshrc.edu.sa or falkuraya@kfshrc.edu.sa
[15]These authors contributed equally to this work.
Received 4 March 2016; revised 2 June 2016; accepted 6 June 2016; published online 19 July 2016

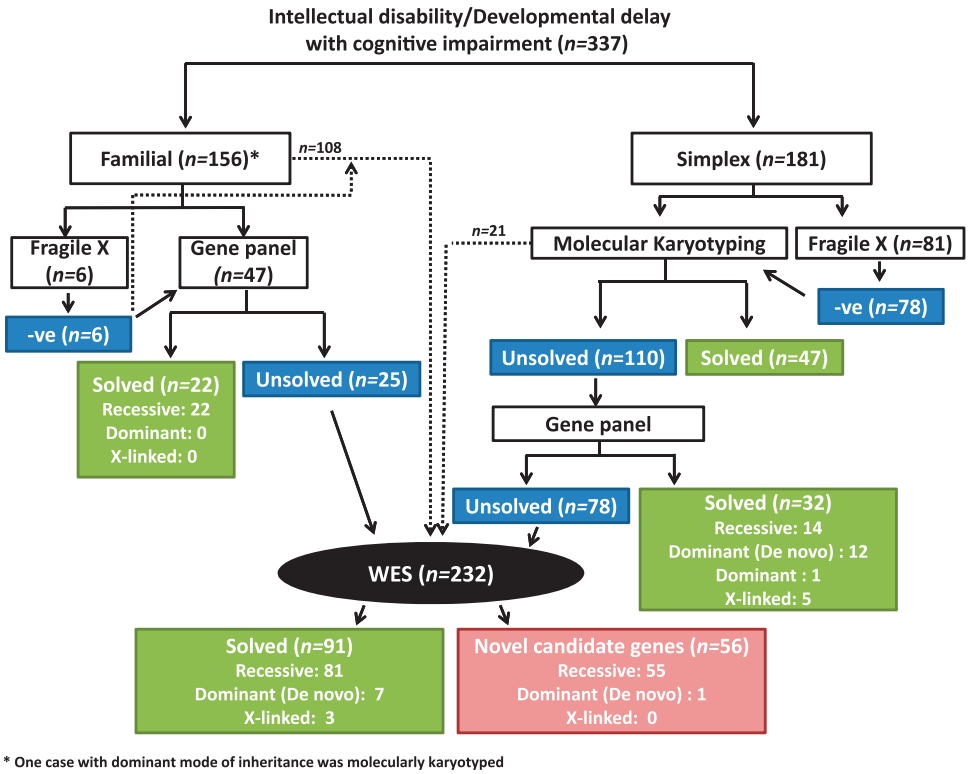


**Figure 1.** Flowchart that summarizes the workflow of the study. WES, whole-exome sequencing.

of genomic tools that are agnostic to the phenotype marked a dramatic change in the diagnostic approach to individuals with ID. Molecular karyotyping greatly improved our understanding of the role of copy number variants in cognitive phenotypes and is now the recommended first-tier diagnostic test in children with ID.[3,4] Genomic sequencing (whole-exome or whole-genome) is a more recent development and has been found to identify causal mutations in up to 50% of cases.[5–8] Thanks to these recent technological developments, the 'morbid genome' of ID, defined as the sum of genetic and genomic mutations identified in the context of ID phenotypes, has expanded greatly in the recent years and is likely to expand further as more patients get tested.

Despite the established diagnostic yield of genomic sequencing in the setting of ID, previously published studies reported on pre-selected individuals, that is, those who had been through the 'routine' testing strategy that failed to identify the causal mutation. Studies that evaluate the clinical diagnostic yield of genomic sequencing on 'naive' ID individuals are needed, however, to inform recommendations regarding the routine use of this technology in clinic. In an attempt to address this gap in our knowledge about the clinical diagnostic yield of genomic sequencing, we report our experience with prospective application of genomic analysis on all individuals with ID (or delayed cognitive development in the case of younger children) who were referred to our clinical genetics service by their pediatricians or pediatric neurologists.

## MATERIALS AND METHODS

### Human subjects

Individuals with a documented intelligent quotient of 70 or less were eligible for the study. Younger children (< 5 years) were eligible if developmental assessment by a pediatric neurologist revealed delayed acquisition of speech and other cognitive developmental domains regardless of whether other developmental domains were also involved (cases were labeled as developmental delay or global developmental delay accordingly). All subjects were evaluated by board-certified neurologists and clinical geneticists. Clinical evaluation included standard medical and family history and clinical examination. Subjects underwent brain imaging (magnetic resonance imaging or computed tomography), a metabolic screen (plasma carnitine, acylcarnitines, amino acids, ammonia and lactate), complete blood count, electrolytes, and liver, renal and thyroid function tests as part of standard clinical evaluation. A written consent was signed by the parents (or legal guardians) of all subjects prior to enrollment (KFSHRC RAC# 2121053). Once enrolled, blood was collected in ethylenediaminetetraacetic acid collection tubes for genetic analysis from the index and available family members in parallel with the standard clinical evaluation.

### Genomic testing algorithm

Simplex cases underwent molecular karyotyping and sequencing by a multi-gene panel that encompasses 758 genes with published link to various neurogenetic diseases as previously described.[9] If negative, we proceeded with whole-exome sequencing (WES). Male simplex cases who lack major facial dysmorphism also underwent FRAXA testing using a standard Southern blot protocol combined with triplet-repeat PCR. Familial cases underwent the multi-gene panel sequencing and, if negative, WES was carried out. Familial cases consistent with dominant inheritance also underwent molecular karyotyping while those consistent with X-linked inheritance also underwent FRAXA testing, in parallel with the multi-gene panel. When negative, WES was pursued (Figure 1). The exception to the above workflow is cases enrolled prior to the availability of the multi-gene panel where this step was replaced by WES directly (Figure 1). Although cases were analyzed for all applicable modes of inheritance, we routinely conducted autozygome analysis for additional support when the identified variants are recessive in nature as described before.[10–12]

The technical details of molecular karyotyping, multi-gene panel and WES, including the rationale of including specific genes in the multi-gene panel, are described elsewhere.[9,13] in brief, molecular karyotyping was performed using CytoScan HD (Affymetrix, Santa Clara, CA, USA). This array platform contains 2.6 million markers for copy number variation (CNV) detection, of which 750 000 are genotyping single-nucleotide polymorphisms and 1.9 million are non-polymorphic probes for the whole-genome coverage. The analysis was performed using the Chromosome Analysis Suite version Cyto 2.0.0.195(r5758). Calling of pathogenic CNVs was in

accordance with the ACMG (American College of Medical Genetics and Genomics) guidelines.[14] The label 'solved' in the context of molecular karyotyping was only used if the CNV met the definition of 'pathogenic' or 'unknown significance-likely pathogenic' according to these guidelines. For the multi-gene panel, design used Ion AmpliSeq Designer software (Life Technologies, Carlsbad, CA, USA). Primers were then synthesized and pooled into two multiplex reactions based upon PCR compatibility minimizing likelihood of primer-primer interactions. For WES, each DNA sample was treated to obtain the Ion Proton AmpliSeq library using Exome Primer Pools, AmpliSeq HiFi mix (Thermo Fisher, Carlsbad, CA, USA). Libraries for the multi-gene panel and WES were run on Ion Proton instrument (Thermo Fisher). Calling of variants in previously reported disease genes using genomic sequencing followed the recently published guidelines by the ACMG.[15] The label 'solved' in the context of single-nucleotide variants (SNVs) was only applied to cases who harbored 'pathogenic' or 'likely pathogenic' SNVs, according to these guidelines, that explain the phenotype.

For novel recessive candidate disease genes (this category is not covered by the ACMG guidelines), we only report those with variants that meet the following criteria: (a) minor allele frequency $< 0.001$ based on 1500 Saudi exomes, (b) fully segregates with the disease by testing all available family members, (c) locus supported by positional mapping data and (d) loss of function (LOF) or at least a likely pathogenic nature of the variants. For novel dominant candidate disease genes (this category is also not covered by the ACMG guidelines), we only report in this study those with variants that meet the following criteria: (a) *de novo* nature of the variant with confirmed paternity, (b) novel based on 1500 Saudi exomes and ExAC and (c) LOF or at least predicted pathogenic nature of the variants based on *in silico* prediction. LOF was defined as nonsense, frameshift indels and canonical splicing mutations. Missense mutations were only considered if at least two of three in silico software (PolyPhen, SIFT and CADD) assigned a high pathogenicity score to the variant (PolyPhen score of $>0.90$, SIFT score $<0.05$ and CADD $>20$). Additional supportive evidence was sought from the published literature (for example, known link to brain development, neuronal function or animal models).

We also conducted our own computational structural analysis of mutants (Supplementary Computational Biology Materials). Sequences were retrieved from the Uniprot database. BLAST and SwissModel[16] were used to search for suitable structural templates in the Protein Data Bank. SwissModel and RaptorX[17] were used to produce homology models. Models were manually inspected, and mutations evaluated, using the Pymol program (pymol.org). Disorder and secondary structure elements were predicted using RaptorX. Transmembrane helices were predicted using Phobius.[18] Functional information was compiled from various resources, including Uniprot, InterPro,[19] and publications associated with the model templates used.

## RESULTS

### Genomic analysis is more sensitive than standard clinical evaluation

The total number of eligible cases was 337. The rate of consanguinity defined as any degree of parental relatedness equal to or closer than third cousins was 76% (255/337). Male to female ratio was 163:173, and 45 and 50% were below 5 years of age, respectively. The proportion of syndromic versus non-syndromic ID was 152:183 (two were equivocal). These and all other characteristics, including relevant clinical data can be found in Supplementary Table S1. All simplex cases were molecular karyotyped ($n=178$) and simplex male ID cases who lacked major dysmorphism as well as familial cases potentially consistent with X-linked inheritance had Fragile-X testing ($n=87$). One or more specific clinical entity was suspected on clinical basis (standard clinical evaluation) in 54 (16% sensitivity) but only 38 were subsequently confirmed by genomic analysis (70% specificity).

On the other hand, the sensitivity of genomic tests (excluding Fragile-X testing) was 57% (193/337) based on previously reported disease genes or CNVs and 74% (249/337) if the variants identified in novel genes are included (see below). Molecular karyotyping revealed pathogenic or likely pathogenic CNVs in 27% of tested cases (48/178). The multi-gene panel revealed a pathogenic ($n=23$) or likely pathogenic ($n=31$) SNVs in 34% of tested cases (54/157). WES uncovered pathogenic or likely pathogenic SNVs in 39% of tested cases (91/232) (Figure 1 and Supplementary Table S1). WES was only applied after a negative multi-gene panel and/or molecular karyotyping in many cases; however, we would like to highlight that in 129 of cases, WES was applied directly because these cases were recruited prior to the availability of the multi-gene panel. This explains the overwhelming majority of cases in which WES identified a mutation in a known disease gene. In six cases, however, we note that the multi-gene panel failed to identify the causal mutation in a known disease gene that was subsequently identified by WES (Supplementary Table S1). The fact that we applied WES directly on 129 of the cases gives us the opportunity to also calculate the diagnostic yield of WES without prior application of multi-gene panel at 60%. A detailed breakdown of the diagnostic yield of the various genomic tests based on age, gender, syndromic vs non-syndromic and consanguinity vs non-consanguinity is provided in Table 1. Of note, although most causal SNVs identified are recessive, the majority of these recessive mutations (65%) were 'private', that is, completely absent in the heterozygous state in 1500 ethnically matched exomes, which is highly consistent with our recent finding that, contrary to conventional assumptions, founder mutations account for a minority of recessive mutation in our population.[20]

### Expanding the morbid genome of ID

*Pathogenic and unknown significance-likely pathogenic CNVs*. Of the CNVs identified in this cohort, eight (15%) are novel (seven were assigned as pathogenic according to ACMG guidelines and one as unknown significance-likely pathogenic), whereas 46 are known pathogenic CNVs. Pathogenic CNVs include *de novo* deletion of 1476 kb (Chr18:47279692-48756541) in 13DG1493, which encompasses *SMAD4* (MIM 600993), and deletion of 502 kb in 15DG1036 (Chr17:44212416-43710395), which encompasses *KANSL1* (MIM 612452), thus confirming the diagnosis of Myhre (MYHRS [MIM 139210]) and Koolen-De Vries syndromes (KDVS [MIM 610443]), respectively, although neither was suspected clinically. Similarly, the *de novo* chr3:70986209-71412654 deletion in 15DG1264 led to complete loss of *FOXP1 (*MIM 605515), which was not suspected clinically despite the overlapping dysmorphology profile with the very few cases that have been reported with *de novo* point mutations this gene.[21] A full list of the identified pathogenic and unknown significance-likely pathogenic CNVs is listed in Supplementary Figure S1 and Supplementary Table S1.

*Pathogenic and likely pathogenic SNVs*

Expanding the allelic spectrum of established disease genes and supporting the candidacy of previously reported candidate genes: Of the 145 pathogenic or likely pathogenic SNVs identified in this study, 68 (47%) are novel and involve previously reported disease genes (Supplementary Table S1, Supplementary Figure S1). The correct clinical diagnosis was only suspected in a minority of the solved cases, partly because of marked phenotypic differences compared with the published phenotype. For example, the deep intronic mutation in *COG5* (confirmed at the real-time polymerase chain reaction level) was associated with global developmental delay, microcephaly, cleft palate, ambiguous genitalia and agenesis of corpus callosum, a constellation that is distinct from the hypotonia, ataxia and cerebellar hypoplasia described in CDG2I.[22] Similarly, Rett syndrome was not suspected in 14DG1903 with a *de novo MECP2* truncating variant because the head circumference remained normal despite the progressive neuroregression. 13DG0035 is another unusual case of phenotypic expansion where a *de novo GNAS* variant was associated with global developmental delay, brain heterotopia, severely hypoplastic scrotum and thyroid agenesis (Table 2).

In addition, we were able to identify additional likely pathogenic alleles that support the candidacy of previously reported candidate disease genes. These include *ASTN1*, *HELZ*,

**Table 1.** Diagnostic yield for each platform based on gender, age, syndromic vs non-syndromic and consanguinity

| *Yield (platforms)* | | *Sex* | | *Age* | | *Consanguinity* | | *Multiplex/simplex* | | *Inheritance* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Total (337)* | *M (163)* | *F (173)* | *< 5 Years (159)* | *> 5 Years (176)* | *Consanguineous (255)* | *Non-consanguineous (80)* | *Multiplex (156)* | *Simplex (181)* | *Dominant* | *Recessive* | *X-linked* | *Non-syndromic ID (183)* | *Syndromic ID (152)* |
| Fragile-X MK | (3/87) 3%<br>(48/178) 27% | (3/87) 3%<br>(21/78) 27% | NA<br>(27/100) 27% | (1/39) 3%<br>(26/90) 29% | (2/48) 4%<br>(22/88) 25% | (1/48) 2%<br>(21/104) 20% | (2/38) 4%<br>(26/72) 36% | (0/6) 0%<br>1/1[a] | (3/81) 4%<br>(47/176) 27% | NA<br>48 | NA<br>0 | NA<br>NA | (3/39) 8%<br>(13/87) 15% | (0/48) 0%<br>(35/88) 40% |
| Gene panel | (54/157) 34% | (22/70) 31% | (31/86) 31% | (21/69) 30% | (31/86) 36% | (43/113) 28% | (11/44) 25% | (22/47) 47% | (32/110) 29% | 13 | 36 | 5 | (30/100) 30% | (23/55) 42% |
| *Exome (with or without gene panel)* | | | | | | | | | | | | | | |
| Known genes | (91/232) 39% | (50/117) 43% | (41/115) 36% | (45/111) 41% | (46/121) 21% | (86/190) 45% | (5/41) 12% | (70/133) 53% | (21/99) 21% | 7 | 81 | 3 | (46/137) 34% | (40/88) 45% |
| (Known genes +candidate genes) | (147/232) 63% | (76/117) 65% | (71/115) 62% | (70/111) 63% | (77/121) 64% | (139/190) 73% | (7/41) 17% | (112/133) 84% | (35/99) 35% | 8 | 136 | 3 | (83/137) 61% | (58/88) 66% |
| *Total diagnostic yield* | | | | | | | | | | | | | | |
| (Known gene/all CNV/Fragile-X) | (196/337) 58% | (97/163) 60% | (99/173) 57% | (93/159) 58% | (101/176) 57% | (151/255) 59% | (44/80) 55% | (93/156) 60% | (97/181) 54% | 68 | 117 | 8 | (92/183) 50% | (103/152) (68%) |
| (Known gene/ candidate genes/ all CNV/Fragile-X) | (252/337) 75% | (122/163) 75% | (129/173) 75% | (118/159) 74% | (132/176) 75% | (204/255) 80% | (46/80) 58% | (135/156) 87% | (117/181) 65% | 69 | 172 | 8 | (129/183) 70% | (122/152) 80% |

Abbreviations: CNV, copy number variations; ID, intellectual disability; NA, not available. [a]Dominant inheritance pattern.

*THOC6*, *WDR45B, ADRA2B* and *CLIP1*, all of which were reported as candidate disease genes based on single mutations.[23–26] The homozygous LOF variant we identified in *C12orf4* in case 16DG0275 is the same we previously published when we reported *C12orf4* as a novel candidate.[25] Both cases have non-syndromic ID (Supplementary Table S1).

Expanding the genetic heterogeneity of ID

Novel genes with two independent homozygous SNVs 12DG1579 who presented with global developmental delay, microcephaly and epilepsy was found to have a homozygous truncating variant in *DENND5A* NM_015213.3:c.3811del: p. (Gln1271Argfs*67). 16DG0219 presented with an identical phenotype and was also found to have a homozygous likely pathogenic variant in the same gene NM_015213.3:c.1622A>G: p. (Asp541Gly) (Table 3, Supplementary Table S1, Supplementary Clinical Data)). *NEMF* was found to harbor a homozygous truncating variant NM_004713.4:c.1235_1236insC: p. (Pro413Serfs*10) in 12DG0891 and her sister who both presented with ID and hypotonia. A homozygous truncating mutation in *DNHD1* (NM_144666.2:c.12347dup: p. (Gln4117Alafs*14) was identified in a case with global developmental delay and cerebellar dysgenesis (16DG0296) (Table 3, Supplementary Table S1, Supplementary Clinical Data). Through an international collaboration, we were able to identify additional patients with overlapping phenotypes and homozygous truncating variants in these two genes (*NEMF*: NM_004713.3:c.2517_2520del: p.Gly841Argfs*27, and *DNHD1*: NM_173589.3: c.103delC: p. (Leu36Trpfs*11), Supplementary Table S1, Supplementary Clinical Data). Comparison of the phenotype of these patients is provided in the Supplementary Clinical Data.

Novel candidate genes Thirty-two genes not previously linked to human diseases were found to have single candidate variants (Table 3, Supplementary Table S1 and Supplementary Clinical Data). These include a *de novo* nonsense mutation in *TADA1*, and homozygous LOF variants in the following novel candidate genes suggesting their complete or near-complete deficiency in the ID subjects who harbor them: *CDH11, PIP5K1A, PIANP, NUDT2, AP3B2, PLK2, QRFPR, UBE4A, PROCA1, TUBAL3, TP53TG5, ATOH1, SLC39A14, BTN3A2, SYDE2* and *ZMYM5*. Previously unreported missense or in-frame variants that are predicted to be pathogenic were identified in the following additional novel genes: *KLHL24*, *MAMDC2*, *USP2*, *C16orf90*, *CPNE6, UFC1*, *HIR, TRERF1, RGL1, FEZF2, ARFGEF3, FAM160B1, SLC45A1, ARHGAP33* and *CAPS2*. Of note, there was a sufficient number of affected members in the family of 10DG0264 that a single locus could be established by positional mapping that spans *NUDT2* (Supplementary Figure S2) providing additional support of pathogenicity. Similarly, positional mapping of the two cases (16DG0295 and 16DG0606) with the candidate variant in *AP3B2* revealed a single shared ROH with the same haplotype (Supplementary Figure S2). A genomic map of the novel variants identified in this study (CNVs, SNVs in known genes and SNVs in novel candidate genes) are shown in Supplementary Figure S1. In addition, 3D modeling data that support the pathogenic nature of missense variants we identified in novel candidate genes are shown in Supplementary Computational Biology Materials.

## DISCUSSION

Several cohorts have been published to describe the diagnostic yield of genomic sequencing in individuals with ID, which ranged from 27 to 50%.[27,28] Those studies clearly demonstrate the usefulness of genomic sequencing compared to molecular karyotyping, which has an average clinical diagnostic yield of 11%.[3] However, because the subjects in those studies are typically pre-selected based on negative 'routine' workup that included sequencing of one or more likely candidate gene, they do not address the question of whether genomic sequencing can be utilized as a first-tier test. This study is an attempt to address this deficiency in the literature.

**Table 2.** Atypical presentations of known disease genes

| *ID code* | *Gene* | *Causal/surviving variant* | *Published phenotype with MIM number on OMIM* | *Atypical features* |
|---|---|---|---|---|
| 12DG0705 | *KIAA0196* | NM_014846.3:c.1669G>A:p.(Ala557Thr) *de novo* | Autosomal-dominant spastic paraplegia-8 (603563) | Infantile onset, ID, no spasticity |
| 13DG0035 | *GNAS* | NM_080425.3:c.2405T>C: p.(Val802Ala) *de novo* | ACTH-independent macronodular adrenal hyperplasia (219080)<br>Osseous heteroplasia, progressive (166350)<br>Pseudohypoparathyroidism Ia (103580)<br>Pseudohypoparathyroidism Ib (603233)<br>Pseudohypoparathyroidism Ic (612462)<br>Pseudopseudohypoparathyroidism (612463) | Metopic craniosynostosis, hydronephrosis, brain heterotopia, thinning of corpus callosum, thyroid agenesis |
| 12DG2577 | *COG5* | NM_001161520.1:c.1120-12T>A homo (confirmed on RT-PCR) | Congenital disorder of glycosylation, type III (613612) | Cleft lip and palate, perforated bowel, underweight, microcephaly severe spasticity, squint, simplified ears and microphthalmus, down-slanting palpebral fissure, bitemporal narrowing, small mouth, absence of corpus callosum, absence of cingulate gyrus and colpocephaly. |
| 11DG1735 | *SIL1* | NM_022464.4:c.1030-9G>A homo (confirmed on RT-PCR) | Marinesco–Sjogren syndrome (248800) | Typical features of Joubert syndrome on MRI |
| 14DG0902 | *SBDS* | NM_016038.2:c.258+2T>C homo | Shwachman–Diamond syndrome (260400)<br>Susceptibility to aplastic anemia (609135) | DD, microcephaly, lack of skeletal, hematological and gastrointestinal features |
| 12DG1367 | *CAPN3* | NM_173087.1:c.1325G>A: p.(Arg442Gln) homo | Muscular dystrophy, limb-girdle, type 2A (253600) | ID, severe hyperlaxity of joints |
| 15DG1372 | *LARS2* | NM_015340.3:c.457A>C: p.(Asn153His) homo | Perrault syndrome 4 (615300) | ID, short stature, T1DM, cirrhosis, thrombocytopenia, leukopenia, seizures, psoriasis, absent ovaries and uterus, vitiligo, osteoporosis |

Abbreviations: ID, intellectual disability; MIM, Mendelian inheritance in man; MRI, magnetic resonance imaging; OMIM, Online Mendelian inheritance in man; RT-PCR, real time-polymerase chain reaction.

Multi-gene panels offer the advantages of relatively low cost and ease of interpretation compared to WES or whole-genome sequencing.[9] By applying this technique in parallel with molecular karyotyping on 226 individuals with ID, we were able to provide a likely molecular diagnosis to 45%, compared with 21% by molecular karyotyping alone. The application of WES to those with negative results on molecular karyotyping and/or multi-gene panel provided a likely etiology in 22% (40% if novel candidate genes are counted). In the hypothetical scenario of having applied WES to all cases, we estimate an overall yield of 43% (60% if novel candidate genes are counted) assuming it will detect all the variants in the multi-gene panel and none of the CNVs, although it is very likely that larger CNVs would also have been identified. Reassuringly, this is consistent with the diagnostic yield we observed when WES was indeed applied in lieu of multi-gene panel before the latter was available. Importantly, we show that the yield (based on known disease genes only) of genomics first approach remains high even if we limit our analysis to non-consanguineous cases (55%), which suggests that our findings have relevance to outbred populations as well.

Consistent with other studies, many of the molecular lesions identified by genomic techniques were not suspected clinically, which highlights their power in overcoming the limited sensitivity and specificity of unaided clinical evaluation of individuals with ID.[29] This new trend of 'reverse phenotyping' or 'genotype to phenotype' made possible by the application of clinical genomics will continue to grow.[30] As shown by the illustrative examples in Table 2, the potential of this approach to unravel the full spectrum of phenotypes associated with each disease gene will greatly enhance our ability to interpret the phenotypic consequences of variants.

One obvious advantage of WES is its ability to identify novel disease genes. As highlighted previously, it is critical that these candidates are made available to facilitate matchmaking, which in turn can establish their bona fide link to disease in humans.[23,25,31] The majority (59%) of the novel candidate genes we report in this study harbor homozygous LOF variants that render the affected individual natural knockout for the respective gene.[32] In the case of *DENND5A*, an additional missense mutation was also identified in a second family with an overlapping phenotype. *DENND5A* encodes a guanine-nucleotide exchange factor that activates Rab39b, which is also mutated in ID patients.[33] Similarly, *NEMF*, which we found to be homozygously truncated in two families with ID, encodes a protein that directly interacts with MECP2 in the brain to form a complex that was proposed to mediate the pathogenesis of *MECP2*, a gene with established link to severe neurodevelopmental disorders in males and females.[34] Although very little is known about the protein encoded by *DNHD1*, we note that this is another gene in which we identified more than one homozygous truncating mutation in two independent families with ID phenotypes, which substantiates the link we propose between *DNHD1* mutations and ID.

Strong links to brain development and function support the candidacy of other candidate genes. For example, *PLK2* deficiency was found to prevent homeostatic shrinkage and loss of dendritic spines, and to impair memory formation, making its biallelic loss of function a likely cause of ID.[35] *QRFPR* is one of four significantly downregulated genes in the prefrontal cortex of the spontaneously hypertensive rat, a model for schizophrenia and attention-deficit/hyperactivity disorder.[36] A knockout mouse model is available for *Ap3b2* and exhibits marked neurobehavioral abnormalities and epilepsy, likely due to abnormal synaptic vesicle protein composition.[37] The pontocerebellar hypoplasia observed in the individual with *ATOH1*-related ID is faithfully recapitulated in the knockout mouse model.[38] Similarly, *CPNE6* is necessary for synaptic plasticity and the knockout mouse displays deficient hippocampal long-term potentiation.[39] The knockout mouse model of *FEZF2* displays abnormal development of the cortex and corticospinal tract.[40–42]

Although there is no available mouse model for *SYDE2*, knockout of its closely related paralog *Syde1* results in reduced docking of synaptic vesicle at the active zone and impaired synaptic transmission.[43] SYDE2 and SYDE1 are the mammalian orthologs of SYD-1, which is required for axonal guidance in Caenorhabditis elegans, and Syd-1, which regulates pre- and postsynaptic maturation in Drosophila.[44,45] *KLHL24* is widely

**Table 3.** List of novel/candidate genes and the corresponding clinical summary

| *ID code* | *Gene* | *Causal/surviving variant* | *Clinical synopsis* | *Supporting evidence* |
|---|---|---|---|---|
| *Novel genes with two independent mutations* | | | | |
| 16DG0219 | *DENND5A* | NM_015213.3:c.1622A>G: p.(Asp541Gly) homo | Early infantile epileptic encephalopathy, blindness, GDD, primary microcephaly, slight dysmorphic features | Mammalian cell- and Xenopus-based studies reported that DENND5A is a regulator of neurite outgrowth during neuronal differentiation (PMID: 26531636), 2nd hit, autozygosity mapping, segregation, no other candidate variants. D541 is located in dDENN domain. Predicted effect: D541G might affect binding and functional efficiency of the interaction of DENND5A and Rab GTPases (see Supplementary materials). |
| 12DG1579 | *DENND5A* | NM_015213.3:c.3811del: p.(Gln1271Argfs*67) homo | Hydrocephalus, holoprosencephaly, seizures, GDD, eczema, asthma, chronic constipation, hypotelorism, abnormality of the cheeks, hypertonia, hyperreflexia | Mammalian cell- and Xenopus-based studies reported that DENND5A is a regulator of neurite outgrowth during neuronal differentiation (PMID:26531636), 2nd hit, LOF, autozygosity mapping, segregation, no other candidate variants |
| 12DG0891 | *NEMF* | NM_004713.4:c.1235_1236insC: p.(Pro413Serfs*10) homo | GDD | Interacts with MECP2 (see text), 2nd hit, LOF, segregation, autozygosity mapping, no other candidate variants. Please see Supplementary Clinical Data for details of the other family with biallelic *NEMF* mutation |
| 16DG0296 | *DNHD1* | NM_144666.2:c.12347dup: p.(Gln4117Alafs*14) homo | GDD, Dandy-Walker malformation, cerebellar dysgenesis. | 2nd hit, LOF, segregation, no other candidate variants. Please see Supplementary Clinical Data for details of the other family with biallelic *DNHD1* mutation |
| *Candidate genes that map to single loci based on linkage analysis in multiplex families* | | | | |
| 10DG0264 | *NUDT2* | NM_001244390.1:c.34C>T: p.(Arg12*) homo | Intellectual disability, hypotonia, normal brain MRI | LOF, positional mapping, segregation, no other candidate variants |
| 16DG0295 | *AP3B2* | NM_001278512.1:c.1837del: p.(Glu613Serfs*182) homo | GDD, seizures, white matter changes on MRI | Suggestive animal model (see text), LOF, positional mapping, segregation, no other candidate variants |
| 16DG0606 | *AP3B2* | NM_001278512.1:c.1837del:p.(Glu613Serfs*182) homo | GDD, seizures, GERD, microcephalic, hypotonic | Suggestive animal model (see text), LOF, positional mapping, segregation, no other candidate variants |
| *Candidate genes with de novo mutations* | | | | |
| 13DG0182 | *TADA1* | NM_053053.3:c.598C>T: p.(Arg200*) | ID, autistic behavior | *De novo*, LOF, no other candidate variants |
| *Candidate genes with predicted loss-of-function mutations* | | | | |
| 11DG1257 | *TUBAL3* | NM_001171864.1:c.316C>T: p.(Arg106*) homo | Microcephaly, cortical dysplasia, intellectual disability, joint laxity, hyperreflexia, strabismus | Expressed in medial prefrontal cortex of rat (PMID: 23376741), LOF, autozygosity mapping, segregation, no other candidate variants |
| 13DG1764 | *PIP5K1A* | NM_001135636.1:c.1078C>T: p.(Arg360*) homo | GDD, GH deficiency, chronic diarrhea, short stature, delayed speech and language development, hypermetropia, abnormality of the head, prominent forehead, horizontal eyebrows, synophrys, deeply set eyes, upslanted palpebral fissures, overbite, abnormality of upper lipEverted lower lip vermilion, hypoplastic nipples, wide intermamillary distance | Brain enriched (PMID: 15018809), LOF, autozygosity mapping, segregation, no other candidate variants |
| 16DG0294 | *UBE4A* | NM_004788.3:c.384G>A: p.(Trp128*) homo | Seizures, microcephaly, obesity, GDD, prominent teeth, small hands, small feet, PWS features | LOF, autozygosity mapping, segregation, no other candidate variants |
| 15DG1898 | *ATOH1* | NM_005172.1:c.212del: p.(Gly71Alafs*36) homo | Generalized hypotonia, nystagmus, poor visual tracking, open mouth, tented upper lip vermilion, hypoplasia of the cerebellum, hypoplasia of the brainstem, pontocerebellar hypoplasia, brain atrophy (frontal lobe) | Suggestive animal model (see text), LOF, autozygosity mapping, segregation, no other candidate variants |
| 14DG0924 | *SLC39A14* | NM_001128431.2:c.313G>T: p.(Glu105*) homo | Neurodegeneration, intellectual disability, hypermanganesemia, abnormal signal of globus pallidus | Transporter of trace elements (see text), LOF, autozygosity mapping, segregation, no other candidate variants |
| 15DG1896 | *TP53TG5* | NM_014477.2:c.255-2A>G homo | Prominent nose, abnormality of the columella, seizures, delayed speech and language development, intellectual disability, highly arched eyebrows, history of hydronephrosis, narrow mouth, short palpebral fissures, macular hyperpigmentation | Brain enriched (PMID: 10719363), LOF, autozygosity mapping, segregation, no other candidate variants |
| 12DG1149 | *CDH11* | NM_001797.3:c.999+1G>T homo | Delayed fine motor development, intellectual disability, delayed eruption of teeth, synophrys, abnormality of the eyebrows, wide nasal bridge, hypertelorism, proptosis, anteverted nares, malar flattening, thin upper lip vermilion, agenesis of incisors, pointed chin, shortening of all phalanges of fingers, wide intermamillary distance, abnormality | Governs the fate of dendritic spine morphogenesis and synaptic functions (see text), LOF, autozygosity mapping, segregation, no other candidate variants |

**Table 3.** (Continued)

| *ID code* | *Gene* | *Causal/surviving variant* | *Clinical synopsis* | *Supporting evidence* |
|---|---|---|---|---|
| | | | of the anus, bifid scrotum, bilateral cryptorchidism, hypospadias. | |
| 14DG0506 | *SYDE2* | NM_032184.1:c.1544C > G: p.(Ser515*) homo | Jaundice, hepatitis, hepatomegaly, microcephaly, GDD, endophthalmitis, seizures, prominent epicanthal folds, synophrys, short nose, hypertonia, abnormality of the corpus callosum, abnormality of the periventricular white matter, delayed myelination, irregular borders of the lateral ventricles, abnormal signal intensity in the right frontal lobe | Regulates axonal guidance (see text), LOF, autozygosity mapping, segregation, no other candidate variants |
| 14DG0268 | *PROCA1* | NM_001304954.1:c.20_21insTC: p.(Ser8Profs*52) homo | Progressive microcephaly, GDD, strabismus, static encephalopathy | Regulates cell cycle, a key mechanism of microcephaly (PMID: 15159402), LOF, autozygosity mapping, segregation, no other candidate variants |
| 15DG0989 | *PIANP* | NM_001244014.1:c.340C > T: p.(Arg114*) homo | Bilateral cryptorchidism, hypotonia, GDD | Enriched in cerebellum (PMID: 21241660), LOF, autozygosity mapping, segregation, no other candidate variants |
| 16DG0032 | *PLK2* | NM_006622.3:c.272del:p.(Gly91Valfs*9) homo | Dysmorphic features, GDD, strabismus | Suggestive animal model (see text), LOF, autozygosity mapping, segregation, no other candidate variants |
| 16DG0121 | *QRFPR* | NM_198179.2:c.373C > T: p.(Gln125*) homo | GDD, ADHD | Suggestive animal model (see text), LOF, autozygosity mapping, segregation, no other candidate variants |
| 16DG0297 | *BTN3A2* | NM_007047.4:c.646G > T: p.(Glu216*) homo | GDD, dysmorphic features | Elevated in the brain of HIV-associated neurocognitive disorders (HAND) (PMID: 26569176), LOF, autozygosity mapping, segregation, no other candidate variants |
| 11DG1531 | *ZMYM5* | NM_001142684.1: c.362_363insT:p.(Asp33Valfs*13) homo | GDD, dysmorphic features, bone lytic lesions, anhidrosis | Regulates the expression of PSEN1, a dementia and Alzheimer disease candidate gene (PubMed: 17126306). Autozygosity mapping, segregation, no other candidate variants |
| *Candidate genes with missense (or in-frame indels) mutations* | | | | |
| 13DG1724 | *CPNE6* | NM_006032.3:c.1193G > A: p.(Arg398His) homo | Neurodegeneration, strabismus, nystagmus, dysarthria, muscle weakness, hyporeflexia, abnormality of the cerebral white matter | Suggestive animal model (see text), autozygosity mapping, segregation, no other candidate variants. R398 is located in vWFA domain. Predicted effect: R398H may affect ligand binding (see Supplementary Computational Biology Materials) |
| 12DG0663 | *KLHL24* | NM_017644.3:c.1609T > A: p.(Cys537Ser) homo | Hypotonia, macrocephaly, nystagmus, myopathic facies malar flattening, smooth philtrum, narrow mouth, high palate, prominent nasal tip, joint laxity, generalized hypotonia, hyporeflexia | KLHL24 is abundantly expressed in the cortex and hippocampus and mediates functional roles of glutamate receptor (see text), autozygosity mapping, segregation, and no other candidate variants. C537 is located in the 5th KELCH domain. Predicted effect: C537S may cause mild tendency for misfolding/loss of structural stability (see Supplementary Computational Biology Materials). |
| 13DG1823 | *USP2* | NM_004205.4:c.550G > A: p.(Gly184Arg) homo | Hypotonia, seizures, developmental delay, cryptorchidism, club feet | Autozygosity mapping, segregation, no other candidate variants. G184 is located in mostly disordered N-terminal region required to bind to MDM4. Predicted effect: G184R might affect subcellular localization and binding to other proteins, in particular MDM2 (see Supplementary Computational Biology Materials). |
| 15DG0585 | *CAPS2* | NM_032606.3:c.1490_1491delinsAA: p.(Gly497Glu) homo | Non-syndromic intellectual disability (NS ID) | Regulates Ca(2+) triggered release of transmitters by modulating synaptic vesicle pool (PMID: 18022372). Autozygosity mapping, segregation, no other candidate variants. G497 is located in the second calcium-binding EF-hand motif. Predicted effect: G497E is predicted to abolish calcium binding to one of the EF-hand motifs. This will impart calcium-sensing and calcium-dependent function of CASP2 (see Supplementary Computational Biology Materials). |
| 15DG0641 | *RGL1* | NM_015149.3:c.2333G > A: p.(Arg778His)<br>NM_015149.3:c.965C > T: p.(Thr322Ile)<br>Compound heterozygous | GDD, seizures, normal brain MRI, no facial dysmorphism | Brain enriched (PMID: 10231032), segregation, no other candidate variants. T322 is located in Ras-GEF domain. Predicted effect: T322I,the non-conservative substitution might affect binding of RGL1 to Ral, possibly weakening the interaction (see Supplementary Computational Biology Materials). |
| 15DG2104 | *FEZF2* | NM_018008.3:c.708_719del: p.(Arg237_Ala240del) homo | GDD, autistic behavior, brain atrophy, microcephaly, suspected DW malformation, dystonia, spasticity, hypotonia, subtle dysmorphic features | Suggestive animal model (see text), autozygosity mapping, segregation, no other candidate variants. |

**Table 3.** (Continued)

| *ID code* | *Gene* | *Causal/surviving variant* | *Clinical synopsis* | *Supporting evidence* |
|---|---|---|---|---|
| 16DG0652 | *ARHGAP33* | NM_052948.3:c.1495G > A:p.(Val499Met) homo | GDD, seizures, microcephaly and dysmorphic features | Autozygosity mapping, segregation, no other candidate variants. ARHGAP33 deficient mice have autism-associated social behaviour and abnormal synapse development (PMID: 26839058). V499 is located in Rho-GAP domain. Predicted effect: V499M may lead to significant destabilization and deformation of the GTPase-binding site (see Supplementary Computational Biology Materials). |
| 12DG0708 | *MAMDC2* | NM_153267.4:c.664T > C: p.(Ser222Pro) homo | Autistic behavior, intellectual disability, brachycephaly, microtia, hypertelorism, short philtrum, thick vermilion border, prominent chin | Autozygosity mapping, segregation, no other candidate variants. S222 is located in the 2nd MAM domain. Predicted effect: S222P might cause mild tendency for altered structural stability and affect ligand interactions (see Supplementary Computational Biology Materials). |
| 14DG0090 | *C16orf90* | NM_001080524.1:c.159G > C: p.(Lys53Asn) homo | Intellectual disability, inguinal hernia, frontal upsweep of hair, macrotia, high palate, hypertonia, hyperreflexia, abnormality of the cerebrum, vitamin D deficiency | Autozygosity mapping, segregation, no other candidate variants. K53N is non-homologous (N is smaller and not charged). It is likely that his protein functions as an adaptor/cofactor and binds to other molecules, possibly proteins. In this case the non-homologous substitution might affect binding (see Supplementary Computational Biology Materials). |
| 13DG1876 | *TRERF1* | NM_033502.2:c.476T > A: p.(Val159Asp) homo | GDD, seizures, hydronephrosis, IUGR, oligohydramnios, upslanted palpebral fissures, unsteady gait, hyperreflexia, abnormality of cerebral white matter, CNS hypomyelination, lactic acidosis | Autozygosity mapping, segregation, no other candidate variants. Predicted effect: The non-conservative substitution V159D might affect binding to protein ligands or DNA (see Supplementary materials). |
| 12DG0178 | *UFC1* | NM_016406.3:c.317C > T: p.(Thr106Ile) homo | FTT, brain atrophy, GDD, hypotonia | Regulates protein turnover, a critical process in the pathogenesis of neurological diseases (PMID: 15071506), segregation, no other candidate variants. Predicted effect: T106I might lead to local distortions that might affect catalytic activity or positioning of UFC1 in a larger complex. Hence this mutation might affect Ufm1-modification (see Supplementary materials). |
| 16DG0088 | *ARFGEF3* | NM_020340.4:c.4541G > A: p.(Arg1514Gln) homo | GDD, hypotonia | Autozygosity mapping, segregation, no other candidate variants. R1514 is likely to be in an importin-like armadillo repeat-motif domain. R1514Q might have a role in allosteric regulation or ligand recognition (see Supplementary Computational Biology Materials). |
| 15DG2696 | *FAM160B1* | NM_001135051.1:c.248T > C: p.(Leu83Pro) homo | ID, dysmorphism | Autozygosity mapping, segregation, no other candidate variants. It can be speculated that the substitution disrupts helices, and hence affects fold and (ligand binding) function (see Supplementary Computational Biology Materials). |
| 15DG2421 | *SLC45A1* | NM_ 001080397.2: c.269T > C : p.(Ile90Thr) homo | GDD, subtle dysmorphism, squint, cerebellar hypoplasia with DW malformation | Brain enriched solute carrier family protein (PubMed: 10729226). Autozygosity mapping, segregation, no other candidate variants. I56 is located in a basic region (KRRKC*I*R) that might interact with charged lipid head groups. Might be involved in ligand interactions (lipid head groups, membrane, but also homo- or heterologous protein interactions) (see Supplementary Computational Biology Materials). |
| 11DG2454 | *HIRA* | NM_003325.3:c.41A > G: p.(Lys14Arg) homo | GDD, severe postnatal growth retardation, strabismus, cryptorchidism, hirsutism (especially over the back with a whorl), horizontal eyebrows, synophrys, wide nasal bridge, long eye lashes, deeply set eyes, strabismus, smooth upper lip vermilion, mandibular prognathia, diastema, short neck, broad nail at index finger, prominent calcaneus, scoliosis. | HIRA transcripts are abundant in the developing neural plate, the neural tube, the neural crest of chicken and regulates chromatin modifications (PMID: 9731536, PMID : 15621527), autozygosity mapping, segregation, no other candidate variants. Predicted effect: K14R might affect ligand binding (see Supplementary Computational Biology Materials). |

Abbreviations: ADHD, attention deficit hyperactivity disorder; CNS, central nervous system; DW, Dandy-Walker; GDD, global developmental delay; GERD, gastroesophageal reflux disease; GH, growth hormone; ID, intellectual disability; LOF, loss of function; MRI, magnetic resonance imaging; PMID, PubMed ID; vWF, von Willebrand factor A.

expressed in rat brain, particularly in the cortex and hippocampus and is involved in glutamate receptors regulation.[46] *CDH11* is also expressed in cortical neurons and has been demonstrate *in vivo* to control migration and differentiation of neuroprogenitors.[47]

Even when the molecular diagnosis is based on a variant in a novel gene with essentially no published data on the natural history of the disease, there is a potential for the molecular diagnosis to influence the clinical management. A good example is our finding of a homozygous truncating mutation in *SLC39A14* in 14DG0924, a girl with unexplained neurodegenerative disease that resulted in progressive dystonia and cognitive impairment with associated lesions in the basal ganglia. *SLC39A14* encodes ZIP14, a transporter of trace elements.[48] Evaluation of trace elements in the affected individual's blood revealed a markedly elevated level of manganese (see Supplementary Clinical Data). This prompted us to initiate chelation therapy with excellent response in terms of manganese level. Clinical monitoring is ongoing.

Genomic testing of individuals with ID offers a higher diagnostic yield than the standard workup. Furthermore, recent studies show that it is cost-effective.[49] The data we present in this study suggest that genomic sequencing should be considered early on in the diagnostic workup of these individuals in parallel with or after a negative result of molecular karyotyping.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

We thank the study families for their enthusiastic participation. This work was supported by KACST Grant 13-BIO1113-20 (FSA) and the Saudi Human Genome Project. The research by STA reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST).

## REFERENCES

1 Maulik PK, Harbour CK. Epidemiology of intellectual disability. In: JH Stone, M Blouin (eds). *International Encyclopedia of Rehabilitation*, 2010. Available at: http://cirrie.buffalo.edu/encyclopedia/en/article/144/.
2 Association AP. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub 2013.
3 Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP *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* 2010; **86**: 749–764.
4 Männik K, Mägi R, Macé A, Cole B, Guyatt AL, Shihab HA *et al.* Copy number variations and cognitive phenotypes in unselected populations. *JAMA* 2015; **313**: 2044–2054.
5 Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 2012; **380**: 1674–1682.
6 de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012; **367**: 1921–1929.
7 Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet* 2016; **17**: 9–18.
8 Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH *et al.* Genome sequencing identifies major causes of severe intellectual disability. *Nature* 2014; **511**: 344–347.
9 Group SM. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol* 2015; **16**: 134.
10 Alkuraya FS. Autozygome decoded. *Genet Med* 2010; **12**: 765–771.
11 Alkuraya FS. The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. *Hum Genet* 2013; **132**: 1197–1211.
12 Alkuraya FS. Discovery of mutations for Mendelian disorders. *Hum Genet* 2016; **135**: 1–9.
13 Al-Qattan SM, Wakil SM, Anazi S, Alazami AM, Patel N, Shaheen R *et al.* The clinical utility of molecular karyotyping for neurocognitive phenotypes in a consanguineous population. *Genet Med* 2014; **17**: 719–725.
14 Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med* 2011; **13**: 680–685.
15 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; **17**: 405–424.
16 Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 2006; **22**: 195–201.
17 Kallberg M, Margaryan G, Wang S, Ma J, Xu J. RaptorX server: a resource for template-based protein structure modeling. *Methods Mol Biol* 2014; **1137**: 17–27.
18 Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 2004; **338**: 1027–1036.
19 Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A *et al.* InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* 2012; **40**(Database issue): D306–D312.
20 Abouelhoda M, Sobahy T, El-Kalioby M, Patel N, Shamseldin H, Monies D *et al.* Clinical genomics can facilitate countrywide estimation of autosomal recessive disease burden. *Genet Med* 2016; e-pub ahead of print.
21 Le Fevre AK, Taylor S, Malek NH, Horn D, Carr CW, Abdul-Rahman OA *et al.* FOXP1 mutations cause intellectual disability and a recognizable phenotype. *Am J Med Genet* 2013; **161**: 3166–3175.
22 Paesold-Burda P, Maag C, Troxler H, Foulquier F, Kleinert P, Schnabel S *et al.* Deficiency in COG5 causes a moderate form of congenital disorders of glycosylation. *Hum Mol Genet* 2009; **18**: 4350–4356.
23 Karaca E, Harel T, Pehlivan D, Jhangiani SN, Gambin T, Akdemir ZC *et al.* Genes that affect brain structure and function identified by rare variant analyses of mendelian neurologic disease. *Neuron* 2015; **88**: 499–513.
24 Beaulieu CL, Huang L, Innes AM, Akimenko M-A, Puffenberger EG, Schwartz C *et al.* Intellectual disability associated with a homozygous missense mutation in THOC6. *Orphanet J Rare Dis* 2013; **8**: 62.
25 Alazami AM, Patel N, Shamseldin HE, Anazi S, Al-Dosari MS, Alzahrani F *et al.* Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep* 2015; **10**: 148–161.
26 Larti F, Kahrizi K, Musante L, Hu H, Papari E, Fattahi Z *et al.* A defect in the CLIP1 gene (CLIP-170) can cause autosomal recessive intellectual disability. *Eur J Hum Genet* 2015; **23**: 331–336.
27 Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y *et al.* Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 2014; **312**: 1870–1879.
28 Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F *et al.* Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014; **312**: 1880–1887.
29 Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, McDonald MT *et al.* Clinical application of exome sequencing in undiagnosed genetic conditions. *J Med Genet* 2012; **49**: 353–361.
30 Alkuraya FS. Natural human knockouts and the era of genotype to phenotype. *Genome Med* 2015; **7**: 48.
31 Shaheen R, Patel N, Shamseldin H, Alzahrani F, Al-Yamany R, ALMoisheer A *et al.* Accelerating matchmaking of novel dysmorphology syndromes through clinical and genomic characterization of a large cohort. *Genet Med* 2016; e-pub ahead of print.
32 Alkuraya FS. Human knockout research: new horizons and opportunities. *Trends Genet* 2015; **31**: 108–115.
33 Han C, Daubaras M, McPherson PS. DENND5A regulates NGF-induced neurite outgrowth in PC12 cells and dendrite patterning of primary hippocampal neurons. *Int J Dev Neurosci* 2015; **47**(Pt A): 116–116.
34 Long SW, Ooi JY, Yau PM, Jones PL. A brain-derived MeCP2 complex supports a role for MeCP2 in RNA processing. *Biosci Rep* 2011; **31**: 333–343.
35 Lee KJ, Lee Y, Rozeboom A, Lee J-Y, Udagawa N, Hoe H-S *et al.* Requirement for Plk2 in orchestrated ras and rap signaling, homeostatic structural plasticity, and memory. *Neuron* 2011; **69**: 957–973.
36 Santoro ML, Santos CM, Ota VK, Gadelha A, Stilhano RS, Diana MC *et al.* Expression profile of neurotransmitter receptor and regulatory genes in the prefrontal cortex of spontaneously hypertensive rats: relevance to neuropsychiatric disorders. *Psychiatry Res* 2014; **219**: 674–679.
37 Seong E, Wainer B, Hughes E, Saunders T, Burmeister M, Faundez V. Genetic analysis of the neuronal and ubiquitous AP-3 adaptor complexes reveals divergent functions in brain. *Mol Biol Cell* 2005; **16**: 128–140.
38 Ben-Arie N, Bellen HJ, Armstrong DL, McCall AE, Gordadze PR, Guo Q *et al.* Math1 is essential for genesis of cerebellar granule neurons. *Nature* 1997; **390**: 169–172.

39 Reinhard J. The function of Copine 6 in the brain. PhD thesis, University of Basel 2012.
40 Molyneaux BJ, Arlotta P, Hirata T, Hibi M, Macklis JD. Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* 2005; **47**: 817–831.
41 Chen J-G, Rašin M-R, Kwan KY, Šestan N. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. *Proc Natl Acad Sci USA* 2005; **102**: 17792–17797.
42 Chen B, Schaevitz LR, McConnell SK. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proc Natl Acad Sci USA* 2005; **102**: 17184–17189.
43 Wentzel C, Sommer JE, Nair R, Stiefvater A, Sibarita J-B, Scheiffele P. mSYD1A, a mammalian synapse-defective-1 protein, regulates synaptogenic signaling and vesicle docking. *Neuron* 2013; **78**: 1012–1023.
44 Hallam SJ, Goncharov A, McEwen J, Baran R, Jin Y. SYD-1, a presynaptic protein with PDZ, C2 and rhoGAP-like domains, specifies axon identity in C. elegans. *Nat Neurosci* 2002; **5**: 1137–1146.
45 Owald D, Khorramshahi O, Gupta VK, Banovic D, Depner H, Fouquet W *et al*. Cooperation of Syd-1 with Neurexin synchronizes pre-with postsynaptic assembly. *Nat Neurosci* 2012; **15**: 1219–1226.
46 Laezza F, Wilding TJ, Sequeira S, Coussen F, Zhang XZ, Hill-Robinson R *et al*. KRIP6: a novel BTB/kelch protein regulating function of kainate receptors. *Mol Cell Neurosci* 2007; **34**: 539–550.
47 Schulte JD, Srikanth M, Das S, Zhang J, Lathia JD, Yin L *et al*. Cadherin-11 regulates motility in normal cortical neural precursors and glioblastoma. *PLoS One* 2013; **8**: e70962.
48 Girijashanker K, He L, Soleimani M, Reed JM, Li H, Liu Z *et al*. Slc39a14 gene encodes ZIP14, a metal/bicarbonate symporter: similarities to the ZIP8 transporter. *Mol Pharmacol* 2008; **73**: 1413–1423.
49 Monroe GR, Frederix GW, Savelberg SM, de Vries TI, Duran KJ, van der Smagt JJ *et al*. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med* 2016; e-pub ahead of print.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)