Tohoku J. Exp. Med., 2017, **243**, 297-309

Next Generation Sequencing and Genome-Wide Genotyping Identify the Genetic Causes of Intellectual Disability in Ten Consanguineous Families from Jordan

Tawfiq Jamal Froukh[1]

[1]Department of Biotechnology and Genetic Engineering-Faculty of Science, Philadelphia University, Amman, Jordan

Intellectual disability (ID), occurs in approximately 1 to 3% of the population and tends to be higher in low-income countries and in inbred communities. Despite the high rates of consanguineous marriages and the likely enrichment for recessive forms of ID, the genetic bases of ID in Jordan are largely unstudied. In this study, whole exome sequencing (WES) and homozygosity mapping were used to identify the genetic causes of ID in ten families from Jordan. The studied families are characterized by consanguineous marriage and having one or more progeny with ID. Likely disease-causing missense mutations were identified in eight families; four families are due to mutations in genes previously implicated with ID and the other four families are due to mutations in genes that are not previously implicated with ID. The novel genes include: BSN (Protein Basson), PTCHD2 (Protein dispatched homolog 3), DHRS3 (Short-chain dehydrogenase/reductase 3), and LGI3 (Leucine-rich repeat LGI family member 3). In addition, copy number variant (CNV) deletion and/or duplication were identified in 2 families; one family with 3.5 mega base (Mb) deletion on chromosome17 previously implicated with Smith Magenis Syndrome, and the other family with a novel combination of deletion and duplication in chromosomes 5 and 11. In this pilot study, four genes and one CNV deletion/duplication are identified for the first time in association with ID. The finding of this study further demonstrates the power of WES and homozygosity mapping for clinical diagnostics of ID in consanguineous families in small populations.

Keywords: epilepsy; heterozygous; IQ; mental retardation; next generation sequencing
Tohoku J. Exp. Med., 2017 December, **243** (4), 297-309.

Introduction

Intellectual disability (ID) is a term used to describe persons who have limited ability to learn. The International Classification of Diseases (ICD10) classified ID into mild, moderate, severe and profound based on intelligence quotient (IQ) levels. Patients with intellectual disability may also have congenital malformations or other neurological features such as epilepsy and autism (Leonard and Wen 2002). ID can be identified in early life stages based on many signs such as motor developmental delay, language delay, memory deficit, difficulties understanding rules, ability to solve problems, and realizing the consequences of their actions (Leonard and Wen 2002).

ID might be caused by environmental factors such as infections, gestational substance abuse, and birth complications. ID can also be caused by Genetic factors such as chromosomal deletions, duplications or other rearrangements of varying sizes, as well as single base insertions, deletions or substitutions (Kaufman et al. 2010). With the latest update on August 8, 2017; SysID (Catalog of ID genes) lists 997 primary ID genes (genes that are confirmed by a sufficient number of patients with mutations and/or by sufficient clinical information) and 641 candidate ID genes (genes that are published but do not yet fulfill the criteria to be in the primary ID gene list) (Kochinke et al. 2016) (http://sysid.cmbi.umcn.nl/). Because not all ID genes have been identified; a definitive genetic diagnosis is still lacking in a large subset of individuals with ID (Mefford et al. 2012).

One powerful approach to discovering novel genes in association with ID is to study populations with high rates of parental consanguinity where recessive forms of disease will be enriched. In Jordan, 39.7% of marriages are consanguineous. The lowest percentage is in the capital Amman where 25.5% of marriages are consanguineous, and the highest in Irbid-north east with 52.1% consanguineous marriages (http://www.consang.net). In Amman, before 1980, the first cousin marriage comprised ~30% of all marriages (Khoury and Massad 1992). Such high percentages

Received September 19, 2017; revised and accepted November 13, 2017. Published online December 19, 2017; doi: 10.1620/tjem.243.297.
Correspondence: Tawfiq Jamal Froukh, Department of Biotechnology and Genetic Engineering, Philadelphia University, Jerash Road, Amman 19392, Jordan.
e-mail: tfroukh@philadelphia.edu.jo

of consanguinity in Jordan result in increased risk for recessive diseases. Among couples with a child genetically diagnosed with a recessive genetic disease in Amman, 69% were offspring of couples who were first degree cousins, compared to 14% from nonconsanguineous marriages (Hamamy et al. 2007).

Modern techniques such as genome-wide genotyping and massively parallel sequencing have proven to be successful in identifying novel genes for Mendelian disorders, specifically ID and particularly in consanguineous populations (Abou Jamra et al. 2011a; de Ligt et al. 2012; Rauch et al. 2012; Yang et al. 2013). In a cohort of consanguineous families from Syria, homozygosity mapping revealed 11 novel loci in 64 with non-specific ID (Abou Jamra et al. 2011b). In an Iranian population, a combination of homozygosity mapping and exome sequencing in 136 consanguineous families with ID revealed: (1) disease-causing mutations in 57% of the studied cases and (2) fifty novel candidate genes (Najmabadi et al. 2011). Likewise, in a population from Jordan, seven consanguineous families with nonspecific ID were investigated using next generation sequencing (NGS) and homozygosity mapping (Reuter et al. 2017). Mutations responsible for ID were identified in three of the seven families, and several other candidates were also revealed. In addition, the genes *DENND5A* (DENN domain-containing protein 5A) and *MBOAT7* (Lysophospholipid acyltransferase 7) were newly implicated in a distinctive subtype of epileptic encephalopathy and ID in studies of consanguineous families (Han et al. 2016; Johansen et al. 2016).

In this study, genome-wide genetic analyses of ten consanguineous Jordanian families using a combination of whole exome sequencing (WES) and homozygosity mapping were performed to further evaluate the genetic basis of ID.

Materials and Methods

Patients

Fifty-six individuals belonging to ten consanguineous families with one or more instances of unexplained intellectual disability (22 individuals with ID) were recruited. The parents in each family are either first, double first, or second cousins (Pedigrees in Figs. 1 and 2). This study was approved by the ethics committee at Philadelphia University-Jordan. Written consents were obtained from the parents of all patients.

The families were visited independently on August 2015 in their homes. Phenotypic data were compiled using three resources: (1) previous medical documents (upon availability), (2) development delay story as supplied by the parents and the relatives, and (3) intel-

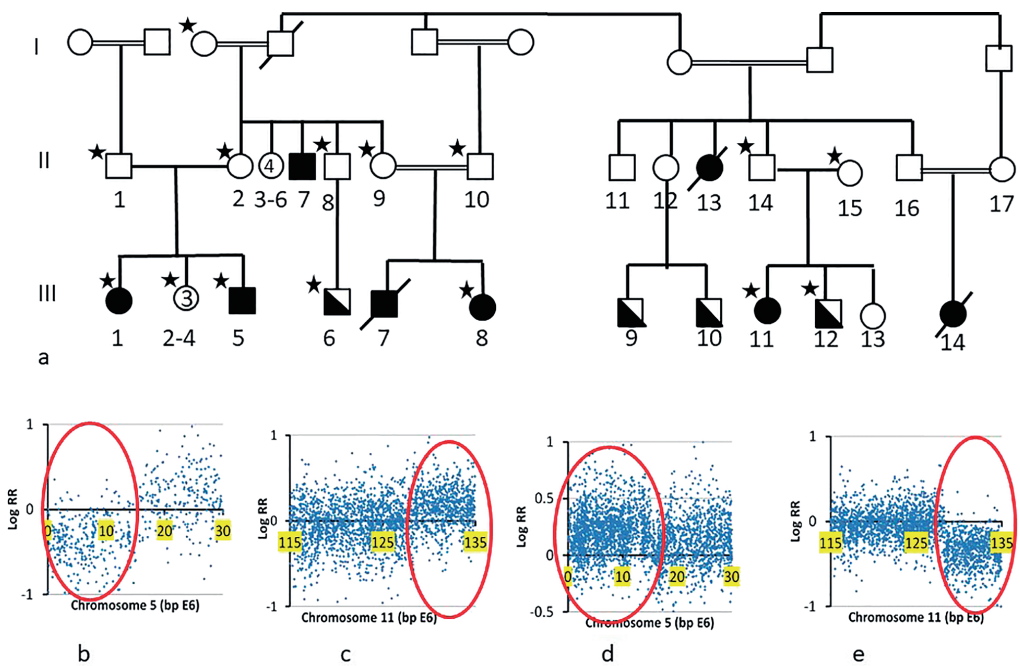

Fig. 1. Family 1 pedigree and the Log RR to chromosome 5 and 11.
(a) Pedigree of the multiplex family 1 showing the 2 groups of affected individuals. The fully shaded squares and circles represent group 1 of individuals with similar phenotypes; while the half-shaded squares and circles represent group 2 of individuals with similar phenotypes (see family 1 in the results and discussion for the details). The star symbol represents the individuals that were recruited and underwent genome-wide genotyping. (b-e) Log RR (log fluorescent intensity value of the probe/baseline value) to chromosome 5 and 11 in the patients of groups 1 and 2. Oval shapes circulate the chromosomal coordinates (measured in base pairs bp×10^6) subjected to deletion or duplication.
(b) Deletion in chromosome 5 detected in individuals III-1, III-5, III-8, and III-11.
(c) Duplication in chromosome 11 detected in individuals III-1, III-5, III-8, and III-11.
(d) Duplication in chromosome 5 detected in individuals III-6 and III-12.
(e) Deletion in chromosome 11 detected in individuals III-6 and III-12.

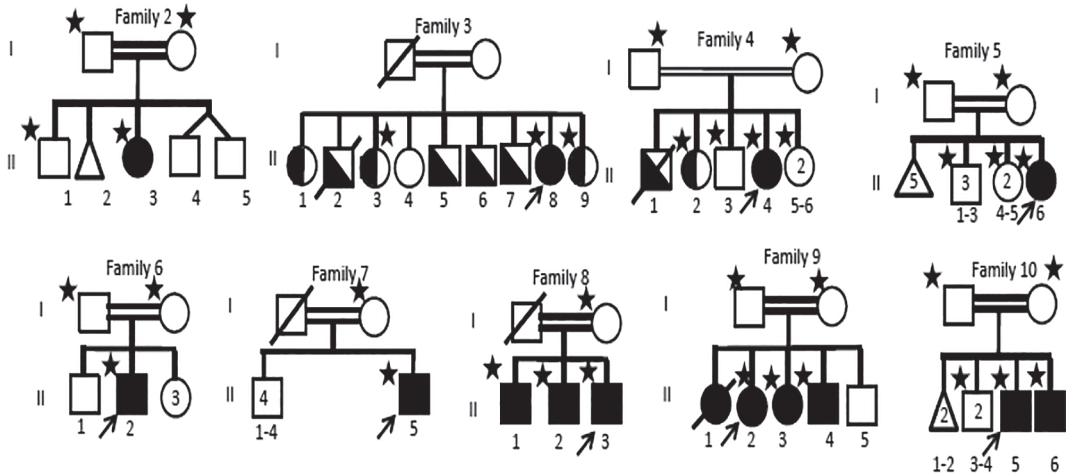

Fig. 2. Pedigrees of the families 2 to 10.
The fully shaded squares and circles represent the probands with ID; while the half-shaded squares and circles represent the individuals with distinctive phenotype other than ID. The star symbol represents the individuals that were recruited and underwent genome-wide genotyping.

ligence level assessment as documented by the center of early diagnosis of disability/ministry of health in Jordan. Phenotypic data were compiled from all families, while only four families were able to provide previous medical documents including (1) magnetic resonance imaging (MRI) reports from three families, (2) Electroencephalogram (EEG) and hearing tests from one family, (3) chromosomal karyotyping and fragile X test from one family, and (4) the results of a vision test from one family.

Peripheral blood was collected from all patients and the available family members. DNA (deoxy ribonucleic acid) was extracted using automated Promega ReliaPrep™ Large Volume HT gDNA Isolation System in combination with HSM 2.0 Instrument.

Genome-wide genotyping array

A total of 56 individuals were genotyped on genome-wide genotyping arrays in the Institute for Genomic Medicine (IGM) at Columbia University Medical Center (CUMC) using standard methods. Two types of BeadChips from Illumnia were used for genotyping: HumanCore-12v1 (50 individuals were genotyped) and OMNI 5 (6 individuals were genotyped). The obtained genotypic data were used: (1) to identify copy number variants (CNV); (2) for mapping of homozygous stretches (SOH) and (3) to identify identity-by-descent (IBD) regions.

CNV

CNV is a term used to refer to the deletion or duplication of DNA segments compared to the reference genome (Wang et al. 2007). Such deletions and duplications were detected from the SNPs (Single Nucleotide Polymorphisms) array using the software PENNCNV (Wang et al. 2007, 2008; Diskin et al. 2008). Screening was for both *de novo* heterozygous CNVs where both parents were available (7 families) and recessive CNVs across all subjects. Families with a likely pathogenic genomic deletion, duplication, or translocation were excluded from further exome sequencing analysis.

Mapping of homozygous stretches

Homozygous regions were calculated using the software Plink v1.90b3.32. Runs of homozygous SNPs were identified with the following settings: (1) window size of 50 SNPs; (2) maximum of five SNPs with a missing call; (3) maximum of two heterozygous called SNPs; and (4) regions of at least 2 mega base (Mb) with a minimum 100 contiguous SNPs (Schuurs-Hoeijmakers et al. 2011). Allelic matching was achieved by segments pairwise comparisons and a match is declared if at least 95% of jointly non-missing homozygous sites are identical. Stretches of homozygosity (SOH) larger than 2 Mb were combined, and the shared stretches of homozygosity (sSOH) between affected individuals within the family were compiled.

IBD regions

IBD regions were calculated using the software Plink with default parameters. IBD regions were used to identify the transmitting parent for inherited heterozygous variants.

WES

The majority of the Mendelian diseases are caused by variants in the coding regions of a gene (exome) which composes around 1% of the human genome (~60 Mb) (Bamshad et al. 2011). Therefore, eight unrelated samples underwent whole exome sequencing and were analyzed for disease-causing point variants and small insertions/deletions.

WES was conducted by including one proband per family. WES was performed using Illumina HiSeq 2500 platform (Illumina) using 64-Mb Roche SeqCap EZ Exome Kit v3.0 (Roche NimbleGen, Madison, WI). Sequencing reads were mapped to the Genome Reference Consortium Human Genome Build 37 (GRCh37) using the program BWA-0.5.10 (Li and Durbin 2010). Polymerase chain reaction duplicates were removed using the program picard-tools-1.59. Single-nucleotide variants and small insertions/deletions (indels) were called using the program GATK-1.6-11 (DePristo et al. 2011) and annotated using the program SnpEff-3.3 (Ensembl-GRCh37.73) (Cingolani et al. 2012). Sequencing was performed in the Genomic Analysis Facility in the IGM at CUMC.

Variants filtration procedure

To identify possible disease-causing mutations, all high-quality variants were identified that are located in the protein coding region

(according to the Ensembl database v68) and/or two base pair flanking splice sites meeting the following quality criteria: (1) at least 10X coverage; (2) quality score (qual) > 30; (3) quality by depth score (QD) > 2; (4) read position rank sum (RPRS) score > –4; (5) mapping quality score > 40; strand bias score (FS) < 200; and mapping quality rank sum score (MQRS) > –6.

The variants meeting the above quality control filters were grouped into the following categories: homozygous variants-recessive model (category 1), compound heterozygous variants-recessive model (category 2) and heterozygous variants-dominant model (category 3). The following filters were applied to generate the candidate variants in each category:

Category 1: Rare homozygous genotypes with minor allele frequency (MAF) ≤ 0.01 were compiled. From this list, the maintained variants are these located: (1) in a SOH for the families where there is only one child with ID, or (2) in a sSOH where there are multiple children with ID, and absent in children with normal intellect. Filtration was then based on predicted effects of the mutation on the protein, maintaining only loss of function (LOF) variants (start lost, stop lost, stop gain, frameshift, splice site acceptor and splice site donor) and nonsynonymous variants predicted to be possibly- or probably-damaging by Polyphen2 Humvar. The putative functional homozygous genotypes in the relevant homozygous stretch were then filtered out if the same recessive genotype was observed in one of the following databases: exome aggregation consortium (ExAC), genome aggregation database (gnomAD), exome variant server (EVS), the greater middle east (GME) Variome Project, or in 4029 IGM-sequenced controls (individuals non-enriched for neuropsychiatric disease and sequenced as part of other genetic studies in the IGM). LOF variants were also excluded if the gene they are located in carry other homozygous LOF variants in EXAC, gnomAD, EVS, GME Variome Project, or in IGM-controls.

Category 2: Candidate compound heterozygous variants were isolated first by limiting to heterozygous variants. The variants where there is a recessive genotype at that site reported in ExAC, gnomAD and EVS, GME Variome Project, or observed in 4029 IGM-sequenced controls were removed. LOF variants were also excluded if the gene they are located in carry other homozygous LOF variants in EXAC, gnomAD, EVS, GME Variome Project, or in IGM-controls. From this list of heterozygous variants, compound heterozygous were isolated by identifying pairs of qualifying variants located in the same gene per proband. IBD was used to identify the heterozygotes variants that were inherited from both parents (true compound heterozygotes) in the five families where genome-wide genotyping on the parents (families 4, 5, 6, 9, 10) was performed. Also, all variants in cases where the parents were not sequenced in close enough proximity to be on the same read were ruled out, and therefore phasing could be inferred by inspecting them using the software Integrative Genomic Viewer (IGV). Candidate compound heterozygous variants were excluded if they were not found in all affected children where DNA was available.

Category 3: Candidate heterozygous variants were first isolated where the alleles are absent in EXAC, gnomAD, EVS, GME Variome Project, or in 4029 IGM-sequenced controls. The maintained variants are the LOF variants in genes that have a probability of LoF intolerance ≥ 0.9, and missense variants predicted to be probably-damaging by Polyphen2 Humvar with a residual variation intolerance score (RVIS) of < 25%. In cases where DNA from the parents was available, Sanger sequencing was performed on the category 3 candidate variants to see if they were de novo variants in the affected child or children.

All variants that passed the filtration steps in the categories 1, 2 and 3 were inspected visually using IGV.

Prioritization scheme for genes reported in the database: online mendelian inheritance in man (OMIM) in association with ID

Recessive variants in categories 1 and 2 were considered likely disease causing if they were in genes reported in OMIM to be: (1) disease associated, (2) recessive inheritance model, and (3) the reported phenotype with the gene reported in OMIM overlapped that of the patient. Candidate disease-causing dominant variants in category 3 were considered likely disease causing if they were in genes reported in OMIM to be disease associated, with dominant inheritance pattern, and the reported phenotype with the gene reported in OMIM overlapped that of the patient.

Prioritization scheme for genes not reported in OMIM in association with ID

In the cases where no disease-causing variant is assigned, the recessive variants in genes not reported in OMIM were investigated to identify the potential genes linked to ID based on: (1) the RVIS, (2) sub-RVIS (sub region RVIS), (3) brain expression pattern (Human Brain Transcriptome Database) and (4) mouse phenotypic overlap (www.informatics.jax.org). RVIS is designed to rank genes in terms of whether they have more or less polymorphic functional variation relative to the given the amount of apparently neutral variation observed. Genes with lower RVIS percentiles have less than expected polymorphic functional variation are considered intolerant (Petrovski et al. 2013). Sub-RVIS is designed to rank exonic regions or protein domains based on their intolerance to functional variation with lower sub-RVIS scores (< 35%) correspond to more intolerant regions located within protein coding exons (Gussow et al. 2016). Human Brain Transcriptome Database provides spatio-temporal information on genes in human brain development. A gene is considered to be brain expressed if the log2 intensity ≥ 6 for a minimum of three periods (de Ligt et al. 2012). Variants are considered to be possible cause of ID if they occur in genes with RVIS percentiles < 25%, subRVIS < 35% (either domain or exon), and either positive brain expression pattern or positive overlap with mouse phenotype.

All assigned variants identified as disease causing were segregated and confirmed by Sanger sequencing.

Results and Discussion

CNV

Diseases causing deletions and/or duplications were detected in two families, which are considered the genetic cause of ID. Family 1 exhibited combination of deletion and duplication in chromosomes 5 and 11 while family 2 exhibited deletion in chromosome 17.

Homozygous mapping

For the detection of SOHs, a size cutoff value of 1Mb is often used in outbred families (Gibson et al. 2006; Curtis et al. 2008; Nalls et al. 2009). Because homozygous mutations in consanguineous families are usually detected in large SOHs reaching several Megabases, the families were analyzed at SOH cutoff value of at least 2 Mb with a mini-

mum 100 contiguous SNPs. The recessive inheritance pattern for the observed phenotypes is hypothesized and therefore sSOHs between the affected siblings and not shared with any of the healthy siblings are identified. The number of SOHs over 2 Mb ranged between 1 and 54 with an average of 23.6 per individual coinciding with the high consanguinity in these families (Schuurs-Hoeijmakers et al. 2011). However, the number of sSOH ranged between 1 and 12 with an average of 5.7 stretches longer than 2 Mb, which represent 24.2% of the individual SOHs per sibling.

WES

The results of WES using Illumina HiSeq 2500 exhibited: (1) an average sequence coverage of 98.9% with 5X base coverage, and (2) an average single nucleotide variant (SNV) of 291290 and an average Indel of 35405. The filtration procedure identified the likely disease-causing variants for 4 families. Three of them are recessive (families 3, 4, 5) and one of them is dominant (family 6). Candidate variants were identified for the families 7, 8, 9 and 10.

Family 1

This is a multiplex family where the parents of the patients III-1 & III-5 and III-11 & III-12 are non-consanguineous (Fig. 1a). The patients III-1, III-5, III-8 and III-11 (group 1; full shaded squares and circles in Fig. 1a) were found to have > 15 Megabase deletion on the p arm of chromosome 5 (5p15.33-p15.1) and > 6 Megabase duplication on the q arm of chromosome 11 (11q23.3-q25) (Fig. 1b, c). While the patients III-6 and III-12 (group 2; half shaded squares and circles in Fig. 1a) show the reciprocal variation with a duplication on the p arm of chromosome 5 (5p15.33-p15.1) and a deletion on the q arm of chromosome 11 (11q23.3-q25) (Fig. 1d, e). ID is the common phenotype among the two groups of the patients, but distinctive phenotypes are observed depending on the two deletion/duplication status. Patients with the chromosome 5 deletion combined to chromosome 11 duplication exhibit motor delay, lack of speech, jerky and unbalanced movement, hyperactive, microcephaly, sleeping disturbance and constipation, whereas, the patients with chromosome 5 duplication combined to chromosome 11 deletion exhibit macrocephaly and speech stammer.

The explanation for the observed two groups is the existence of chromosomal translocation in one of the parents per each patient between the segments 5p15.33-p15.1 and 11q23.3-q25. Such parents do not exhibit abnormal phenotypes because they are not losing genetic material. However, chromosomal malsegregation during the gametogenesis is likely to occur leading to gametes with 5p15.33-p15.1 deletion and 11q23.3-q25 duplication as observed in the patients of group1 or 5p15.33-p15.1 duplication and 11q23.3-q25 deletion as observed in the patients of group 2. The possible explanation for the deceased offspring II-13, III-7 and III-14 is the gametes resulted from chromosomal malsegregation which produces almost the entire chromosome 5 or 11 deletion/duplication. Furthermore, 10 individuals out of 14 (71%) in the third generation are affected and presumably have a translocation. This excess above the expected 50% for offspring of non-consanguineous with a single parent carrying the translocation could be due to consanguinity in the parent's ancestors resulting in both parents carrying the translocation.

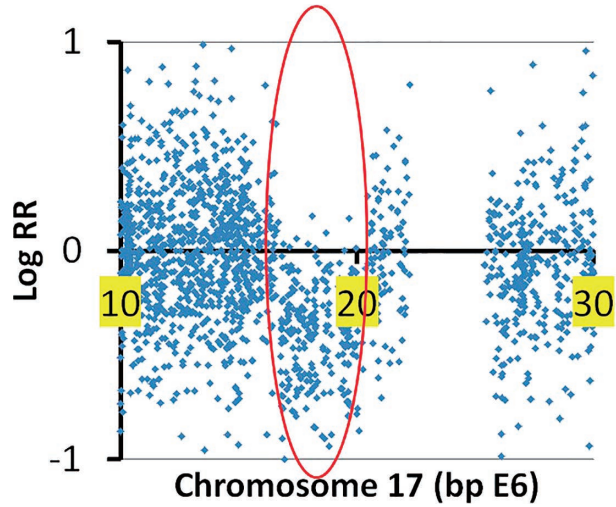

Fig. 3. Log RR to chromosome 17 in family 2.
Log RR (log fluorescent intensity value of the probe/baseline value) to chromosome 17 in the proband of family 2. Oval shapes circulate the chromosomal coordinates (measured in base pairs bp×10^6) subjected to deletion.

Family 2

The parents are first cousins with four living offspring and one spontaneously aborted fetus. The proband 12 years old girl has ID, motor delay, aggressive, ichthyosiform skin, recurrent infection (ear & lung), cataracts, thick eyebrow and short palms and feet. All other siblings were reported by the parents to be healthy. A CNV deletion was detected in the patient that was not detectable in the parents. The deletion size is > 3.5Mb on chromosome 17p11.2 (Fig. 3). The deletion originates a stretch of chromosome inherited from the father. Deletions in this region are the cause of Smith Magenis Syndrome (SMS; Table 1) (Smith et al. 1986).

Family 3

The phenotypes of the affected individuals in this family are heterogeneous. While only one is intellectually disabled and blind (II-8; 14 years old) the others are not intellectually disabled but either: paralyzed (II-9) or have a combination of paralysis, immature growth and blindness (II-3, II-5, II-6 and II-7) (pedigree in Fig. 2).

Individual II-8 with ID was sequenced and the filtration revealed 12 variants in category 1 and 7 variants in category 3 (Table 2). From the 12 variants in category 1 two variants are in genes reported with OMIM diseases but one

Table 1. Clinical features match between the proband in family 2 with the recorded in OMIM.

Category	Smith Magenis Syndrome OMIM#182290 Clinical features	Matching
Head	Brachycephaly	Yes
Face	Midface hypoplasia	No
	Broad face	Yes
Ears	Hearing loss (conductive and/or sensorineural)	No
Nose	Broad nasal bridge	Yes
Heart	Congenital heart defect	Unknown
Kidneys	Structural renal anomalies	Unknown
Spine	Scoliosis	Unknown
Hands	Brachydactyly	No
Central Nervous System	Speech delay	Yes
	Mental retardation (IQ 20-78)	Yes
	Sleep disturbance	No
	Structural brain abnormalities	Unknown
Peripheral Nervous System	Peripheral neuropathy	Yes
	Decreased pain sensitivity	Yes
	Normal nerve conduction velocities	Unknown
	Decrease/absent deep tendon reflexes	Unknown
Behavioral Psychiatric Manifestations	Hyperactivity	Yes
	Polyembolokoilamania (insertion of foreign bodies into body orifices)	Yes
	Behavioral problems	Yes
	Self-destructive behavior	Yes
	Onychotillomania (pulling out nails)	Yes
	Wrist-biting	No
	Head-banging	Yes

Table 2. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 3.

Inheritance	Category	Variants filtered in step 3		OMIM disease gene	Compatible inheritance pattern	OMIM Phenotype overlaps patient phenotype
		Gene	annotation			
recessive	Cat.1	FAM53A	Missense; chr4-1657027-G-A	X		
		ENPEP	Missense; chr4-111441428-C-G	X		
		CHD4	Missense; chr12-6707169-T-G	X		
		YBX3	Missense; chr12-10862594-C-G	X		
		PRB4	Missense; chr12-11461786-C-G	X		
		KIF21A	Missense; chr12-39716603-C-T	√	X	
		GPR84	Stop gained; chr12-54756984-G-A	X		
		PLD2	Missense; chr17-4721857-C-T	X		
		GUCY2D	Missense; chr17-7916436-C-T	√	√	√
		SMYD4	splice site donor; chr17-1703275-AGAG-A	X		
		SCIMP	Missense; chr17-5138028-AC-A	X		
		ZNF555	Missense; chr19-2853649-C-T	X		
Dominant/ Cat.3		ZFP62	Missense; chr5-180275991-G-A	X		
		COL12A1	Missense; chr6-75862138-C-T	√	X	
		ABCC1	Missense; chr16-16138452-A-G	X		
		NOX4	Missense; chr11-89088208-C-T	X		
		DMXL2	Missense; chr15-51751856-G-C	√	X	
		KLHL14	Missense; chr18-30260401-G-C	X		
		BEST2	Missense; chr19-12866534-C-G	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

of them NM_000180.3: c.2129C>T;p.(Ala710Val) in the gene GUCY2D (retinal guanylyl cyclase 1) is reported with an autosomal recessive inheritance pattern with overlapped phenotypes. The phenotypes of the affected individual II-8 show significant overlap with Leber congenital amaurosis 1 (Perrault et al. 1996) (LCA1; OMIM#204000), including mental retardation, central vision loss, blindness, cataract, nystagmus and retarded growth (Table 3). Based on these findings, and no clear support for any of the other candidate variants, the variant in *GUCY2D* was deemed to be likely cause of ID for the proband II-8 this family.

Family 4

In this family, first cousin parents have two affected daughters, one is intellectually disabled (II-4) and the other with occasional epilepsy (II-2), and a deceased son with epilepsy. Epilepsy exists as well in two cousins of the probands and in one of their uncles pedigree in Fig. 2. The proband that was sequenced is II-4 (7 years by the time of examination) who is also characterized by ID, motor delay, coarse face, microcephaly, hyperactive, aggressive, hearing loss, congenital hip dislocation, sleeping disturbance, and abnormal EEG record.

Table 3. Clinical features match between the proband in family 3 with the recorded in OMIM.

Category	Leber congenital amaurosis 1 OMIM#204000 Clinical features	Matching
Growth	Retarded	Yes
Ears	Sensory hearing loss	No
Eyes	Pigmentary retinopathy	No
	Central vision loss	Yes
	Blindness	Yes
	Fundus atrophy	Unknown
	Cataract	Yes
	Keratoconus	Unknown
	Photophobia	Unknown
	Eye poking	Unknown
	Nystagmus	Yes
	Reduced electroretinogram	Unknown
Liver	Hepatomegaly	Unknown
Central Nervous System	Mental retardation	Yes

Table 4. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 4.

Inheritance	Category	Variants filtered in step 3		OMIM disease gene	Compatible inheritance pattern	OMIM Phenotype overlaps patient phenotype
		Gene	Annotation			
recessive	Cat.1	LBR	Missense; chr1-225592109-A-T	√	√	X
		PPP2R1B	Missense; chr11-111630667-C-T	√	√	X
		CD3E	Missense; chr11-118184539-C-T	√	√	X
		KRT34	Missense; chr17-39535724-C-G	X		
		NAGLU	Missense; chr17-40693137-G-A	√	√	√
Dominant/Cat.3		PCDH10	Missense; chr4-134073369-G-A	√	X	
		FAM47E	Missense; chr4-77177671-A-G	X		
		ASH2L	Splice site donor; chr8-37963258-T-G	X		
		CPSF1	Missense; chr8-145621703-C-A	X		
		FBXO31	Missense; chr16-87369816-G-C	√	X	
		CACNA1A	Missense; chr19-13338253-C-T	√	√	X
		MCM3	Missense; chr6-52141973-G-C	X		
		GLI1	Missense; chr12-57859599-G-T	X		
		FLT1	Missense; chr13-28893603-G-T	X		
		CD19	Missense; chr16-28950223-C-A	X		
		RTTN	Missense; chr18-67788926-C-T	X		
		SCL271	Missense; chr19-17611324-G-A	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

A total of 17 candidate variants were identified in categories 1 and 3, seven of them are in genes associated with OMIM diseases (Table 4). Among these was a homozygous variant in the gene *NAGLU* (alpha-N-acetylglucosaminidase) NM_000263.3:c.934G>A;p.(Asp312Asn) a gene associated with Mucopolysaccharidosis type IIIB (Sanfilippo B; OMIM#252920) (Chinen et al. 2005). The variant is heterozygous with the wild type allele in the parents and in her siblings. There is significant phenotypic overlap including mental retardation, sleep disturbances, slower mental development, hyperactivity and aggressive behavior, hearing loss, coarse face, coarse hair, clear corneas and joint stiffness (Table 5). Given the phenotypic similarity between the proband II-4 and that previously reported in the literature, this variant was deemed to be responsible for ID in the proband II-4 this family.

Family 5

First cousin parents experienced five spontaneous abortions, five healthy children, and one intellectually disabled daughter (II-6; 11 years by the time of examination). The proband has an intellectually disabled cousin, also the product of a consanguineous union, with a similar phenotypic presentation. Both children have developmental and speech delay, brain atrophy (based on CT-scan) and macrocephaly. The patient studied in this study (II-6) had a normal karyotype and was negative for fragile X-syndrome.

A total of 15 variants were filtered by from the three categories, four of them are in genes associated with OMIM diseases from which two are with the correct inheritance pattern (Table 6). Among these was a homozygous variant NM_138425.2:c.33G>C;p.Leu11Phe in the gene C12orf57 (protein C10) responsible for Temtamy syndrome (OMIM#615140) (Akizu et al. 2013). The variant is heterozygous with the wild type allele in the parents and in all siblings. This affected child has clinical features consistent with Temtamy syndrome, including mental retardation, autistic features, hypotonia and global developmental delay in the central nervous system, in addition to long face, low-set ears and pes planus in the feet (Table 7). DNA was not available from the affected cousin.

Family 6

An eleven-year-old boy was premature (6 months) and has ID, microcephaly, spastic cerebral palsy, one-time seizure, divergent squint, vertical strabismus and mild hydro-

Table 5. Clinical features match between the proband in family 4 with the recorded in OMIM.

Category	Mucopolysaccharidosis type IIIB (Sanfilippo B) OMIM#252920 Clinical features	Matching
face	Coarse	Yes
Ears	Hearing loss	Yes
Eyes	Clear corneas	Yes
	Synophrys	Unknown
Heart	Asymmetric septal hypertrophy	Unknown
	Cardiomegaly	Unknown
Nasopharynx	Frequent upper respiratory infection	No
Ribs	Thickened ribs	No
Liver	Mild hepatomegaly	Unknown
spleen	Mild splenomegaly	Unknown
Gastrointestinal	Diarrhea	Unknown
Skull	Dense calvaria	Unknown
spine	Ovoid thoracolumbar vertebrae	Unknown
limbs	Mild joint stiffness	Yes
Hair	Coarse	Yes
Central Nervous System	Mental retardation	Yes
	Seizures	No, abnormal EEG
	Sleep disturbances	Yes
	Neurologic deterioration, progressive	No
	Slowing mental development 1.5-3 years	Yes
	Hyperactivity	Yes
	Aggressive	Yes
	Severe behavioral problems age 3-4 years	Yes

Table 6. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 5.

Inheritance	Category	Variants filtered in step 3		OMIM disease gene	Compatible inheritance pattern	OMIM Phenotype overlaps patient phenotype
		Gene	annotation			
recessive	Cat.1	MYLK	Missense; chr3-123427694-T-C	√	X	
		VCPIP1	Missense; chr8-67578388-A-C	X		
		PPFIBP2	Missense; chr11-7654123-C-G	X		
		C12orf57	Missense; chr12-7053317-G-C	√	√	√
		MT-CO3	Missense; chrMT-9229-A-T	X		
	Cat.2	EYS	Missense; chr6-64776324-G-A/ chr6-65300769-G-A	√	√	X
Dominant/Cat.3		GABRG1	Missense; chr4-46067494-A-C	X		
		ING3	Missense; chr7-120607673-C-T	X		
		FMNL3	Missense; chr12-50055767-A-G	X		
		UNC79	Missense; chr14-94120304-C-T	X		
		KPNB1	Missense; chr17-45740491-A-C	X		
		ZNF837	Missense; chr19-58880578-C-T	X		
		FAM135A	Missense; chr6-71234655-A-G	X		
		KIAA1217	Missense; chr10-24762173-T-A	X		
		MMP2	Missense; chr16-55523718-G-A	√	X	

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

Table 7. Clinical features match between the proband in family 5 with the recorded in OMIM.

Category	Temtamy syndrome OMIM#615140 Clinical features	Matching
face	Long face,long philtrum and frontal bossing	Yes
Ears	Low set ears	yes
Eyes	Hypertelorism	no
Nose	Beaked nose	no
Teeth	Dental crowding	no
Pelvis	Hip dislocation	no
Limbs	Genua valgum	no
Hands	Brachydactyly	no
Feet	Pes planus	Yes
Central nervous system	Global developmental delay	Yes
	Mental Retardation	Yes
	Hypotonia	Yes
	Seizure	no
	Corpus callosum abnormalities	unknown
	Autistic features	Yes

cephalus manifested by triventricular dilation. His parents are first cousins and no history for ID in the family. A total of 18 variants in the three categories were identified, including three in genes associated with OMIM disease in the expected inheritance pattern (Table 8). Among these was the heterozygous variant NM_006086.3:c.136C>T; p.(Arg46Trp) in the gene TUBB3 (tubulin beta-3 chain) associated with cortical dysplasia complex with other brain malformations (Table 9; OMIM#614039) (Poirier et al. 2010). The patient presents with a concordant phenotype to

Table 8. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 6.

Inheritance	Category	Variants filtered in step 3		Reported with OMIM disease	Compatible inheritance pattern	OMIM Phenotype overlaps patient phenotype
		Gene	annotation			
recessive	Cat.1	NR1I2	Missense; chr3-119526275-G-A	X		
		PLXNA1	Missense; chr3-126752807-C-T	X		
	Cat.2	OBSCN	Missense; chr1-228467732-G-T/ chr1-228554544-C-T	X		
		CAPN9	Splice site acceptor/missense chr1-230925937-T-G/ chr1-230928608-C-T	X		
		CDH23	Missense; chr10-73472463-G-A/ chr10-73491879-C-T	√	√	X
		KRTAP5-5	Missense; chr11-1651239-G-A/ chr11-1651496-T-G	X		
		JAK3	Missense/frameshift chr19-17941341-A-G/ chr19-17950343-GC-G	X		
Dominant/Cat.3		CYR61	Missense; chr1-86047744-A-T	X		
		HDLBP	Missense; chr2-242186295-G-C	X		
		ARCN1	Frameshift; chr11-118472430-TC-T	X		
		TUBB3	Missense; chr16-89999057-C-T	√	√	√
		UNK	Splice site acceptor; chr17-73818640-A-C	X		
		JAK3	Frameshift; chr19-17950343-GC-G	√	X	
		UTRN	Missense; chr6-145157543-G-T	X		
		FANCI	Missense; chr15-89857883-T-C	X		
		MYLK3	Missense; chr16-46774054-T-A	X		
		FAM65A	Missense; chr16-67572937-G-A	X		
		TANC2	Missense; chr17-61278178-G-A	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

Table 9. Clinical features match between the proband in family 6 with the recorded in OMIM.

Category	Cortical dysplasia complex with other brain malformations OMIM#614039	Matching
	Clinical features	
Head	Microcephaly	Yes
Eyes	Strabismus and nystagmus	Yes
Central nervous system	Mental retardation	Yes
	Delayed psychomotor development	Yes
	Delayed speech language	Yes
	Seizure	Yes
	spasticity	Yes

Table 10. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 7.

Inheritance	Category	Variants filtered in step 3		Reported with OMIM disease	Compatible inheritance pattern	OMIM Phenotype overlaps patient phenotype
		Gene	annotation			
recessive	Cat.1	NDRG2	Missense; chr14-21488690-T-A	X		
	Cat.2	TTN	Missense; chr2-179414964-T-A/ chr2-179613637-A-G	√	√	X
		BSN	Missense; chr3-49700579-G-T/ chr3-49701080-G-A	X		
		MROH2B	Missense; chr5-40999820-C-T/ chr5-41038921-G-A	X		
Dominant/Cat.3		KIAA1109	Missense; chr4-123165036-A-C	X		
		PCDH17	Missense; chr13-58208898-T-C	X		
		CYP19A1	Missense; chr15-51534969-T-C	√	√	X

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

that previously reported for patients with *TUBB3* mutations. Sanger sequencing of the proband and parent confirmed that the variant is *de novo* in the proband. The unaffected children were also found to be wildtype at this locus.

Family 7

The proband is 6-years old boy by the time of examination and his parents are first cousins with no previous history for ID. His clinical features include ID, macrocephaly, astigmatism, short stature, flat feet, kypholordosis, strabismus, prominent CSF space bifrontally & both temporal lobes, anorchia, small scarred right kidney with low function, breath crisis in winter and sluggishness. The clinical geneticist suspected in Prader-Willi syndrome; however, the CNV analysis did not identify deletions in the chromosomal segment (15q11.2) responsible for Prader-Willi syndrome.

A total of 7 variants in the three categories were filtered, two in genes reported in OMIM with a compatible inheritance pattern but the phenotypes were inconsistent with known disease phenotypes (Table 10). Therefore, the recessive variants in genes not reported in OMIM (3 genes) were investigated. Based on the RVIS, sub_RVIS, brain expression and the mouse phenotype in MGI the compound heterozygote variants NM_003458.3:c.10988G>T;p.Gly3663Val and NM_003458.3:c.11489G>A;p.Gly3830Asp in the gene BSN (protein Basson) are highlighted as a possible causative to ID of the proband II-5 in this family (Table 11).

Family 8

First cousin parents have three boys characterized by mild ID and epilepsy. The deceased father has a sibling and

Table 11. Prioritization scheme for genes not reported in OMIM in association with ID in the families 7 to 10.

Family	Gene	0.05%_anypopn_RVIS [ExAC]	subRVIS% Domain	subRVIS %Exon	Brain expressed	Abnormal mouse phenotype (nervous, skeleton, reproductive & renal/urinary, growth/size/body, vision) resulted from targeted mutation /MGI
7	NDRG2	0.058196	54.6	50.1	Yes	Decrease neuron apoptosis & astrocyte number, abnormal: glial cell & astrocyte physiology, & microglial cell chemotaxis
	IL25	0.418714	43.4	75.5	No	No
	BSN	-3.60733	0.018/0.018	68.4/68.4	Yes	No
	MROH2B	2.528691	99.8/99.8	22.74	No	No
8	PTCHD2	-0.14568	2.07477	90.7681	Yes	No
	TNFRSF8	-0.11504	44.0608	76.9078	No	No
	CROCC	5.124837	99.9826	NA	No	Retinal photoreceptor degeneration
	EXTL1	0.898569	98.0656	91.5577	Yes	No
	MT-CYB	NA	NA	NA	NA	No
	BGLAP	0.556493	77/94	76/91	NA	No
	GSE1	-0.57348	10/10	25/73	NA	No
9	DHRS3	0.615842	92	94	Yes	No
10	LGI3	-0.436	27	2	Yes	No

Table 12. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 8.

Inheritance	Category	Variants filtered in step 3		Reported with OMIM disease	Compatible inheritance pattern	Matching with the probands II-1, II-2 and II-3
		Gene	annotation			
recessive	Cat.1	PTCHD2	Missense; chr1-11595087-C-T	X		
		TNFRSF8	Missense; chr1-12175684-G-A	X		
		CROCC	Missense; chr1-17298083-C-T	X		
		EXTL1	Missense; chr1-26349888-C-T	X		
		MT-CYB	Missense; chrMT-15327-C-T	X		
Dominant/Cat.3		RAB3GAP	Stop gained; chr1-220384652-G-A	√	X	
		CAD	Missense; chr2-27459279-G-A	√	X	
		SCRIB	Missense; chr8-144887419-G-A	X		
		PLCG2	Missense; chr16-81953171-C-A	√	√	X
		MCF2L2	Missense; chr3-182923741-G-A	X		
		HIP1R	Missense; chr12-123345236-G-A	X		
		UBN1	Missense; chr16-4926942-C-G	X		
		GSE1	Missense; chr16-85704603-C-G	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

an uncle with epilepsy. The proband II-1 is 8 years old and characterized by motor developmental delay and seizure frequency once every 7 months. The proband II-2 is 6 years old exhibited only one seizure as he was 5 years old. The proband II-3 is 4 years old with motor developmental delay, cannot walk, say very few words and has 18 seizures within 3.8 months. The proband II-3 was sequenced and the filtration yields a total of 13 variants in the two categories 1 and 3 and no compound heterozygous variants (cat. 2) were identified (Table 12). Three variants are in genes reported in OMIM including one with a consistent inheritance pattern but the phenotypes are not matching. Reviewing the variants in genes not reported with diseases in OMIM highlights the variant NM_020780.1:c.3555C>T;p.Arg72Trp in the gene PTCHD2 (protein dispatched homolog 3) as a possible causative (Table 11).

Family 9

Two girls and their brother with ID, motor delay belong to a first cousin parents and have two cousins with similar phenotypes. The proband that was sequenced is II-2 (16 years by the time of examination) who is distinctively aggressive. While a total of 8 variants from the three categories were filtered, three variants are in genes reported in OMIM, one of them is consistent with the inheritance pattern but without overlapping phenotypes (Table 13). Reviewing the variants in the genes not reported in OMIM highlights the variant NM_004753.4:c.328G>A;pVal110 Ile in the gene DHRS3 (short-chain dehydrogenase/reductase 3) as a possible causative (Table 11).

Family 10

Two affected boys (22 and 19 years old by the examination time) belong to a first cousin parents are characterized by ID, motor and speech delay, dolichocephaly (by childhood), aggressive (by childhood), constipation and a total of 3-4 times seizure since born. The proband II-5 was sequenced and the filtration yields a total of 10 variants in the three categories where none of them is a likely disease causing (Table 14). Reviewing the recessive variants in genes not previously implicated with OMIM disease revealed the variant NM_139278.2:c.991G>A;p.Asp331Asn in the gene LGI3 (leucine-rich repeat LGI family member 3) as a possible causative to ID (Table 11). This variant is homozygous in the two probands and heterozygous with the wild type allele in the parents and in one of the healthy sib-

Table 13. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 9.

Inheritance	Category	Variants filtered in step 3		Reported with OMIM disease	Compatible inheritance pattern	Matching with the probands II-2, II-3 and II-4
		Gene	annotation			
recessive	Cat.1	DHRS3	Missense; chr1-12640562-C-T	X		
		FUZ	Missense; chr19-50312016-C-A	√	X	
	Cat.2	TTN	Missense; chr2-179419345-T-C/ chr2-179610985-C-T	√	√	X
Dominant/Cat.3		SASS6	Missense; chr1-100573438-T-G	√	X	
		EBF2	Missense; chr8-25702147-C-G	X		
		NFIB	Missense; chr9-14088311-G-T	X		
		RELN	Missense; chr7-103155887-A-T	X		
		A1CF	Missense; chr10-52580396-C-A	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

Table 14. Variants that passed the filtration steps in the categories 1, 2 and 3 in family 10.

Inheritance	Category	Variants filtered in step 3		Reported with OMIM disease	Compatible inheritance pattern	Matching with the probands II-5 and II-6
		Gene	annotation			
recessive	Cat.1	LGI3	Missense; chr8-22006329-C-T	X		
	Cat.2	TTN	Missense; chr2-179497417-C-T/ chr2-179549474-G-A	√	√	X
Dominant/Cat.3		ADCY2	Missense; chr5-7757649-C-T	X		
		PLEKHA7	Missense; chr11-16804364-G-A	X		
		FADS2	Missense; chr11-61608096-G-A	X		
		AFG3L2	Missense; chr18-12366978-G-C	√	X	
		CACNA1A	Missense; chr19-13409764-C-T	√	√	X
		MYH6	Missense; chr14-23854145-C-G	√	X	
		INTS2	Missense; chr17-59981887-G-A	X		
		ZNF569	Missense; chr19-37904419-G-C	X		

Genes are tagged by √ or X based on the compatible criteria from the OMIM database. Shaded cells represent a nonapplicable criteria.

lings, and homozygous to the wild type allele in the other healthy sibling.

Conclusion

Exome sequencing and homozygosity mapping are used to diagnose the genetic causes in ten consanguineous families with intellectually disabled progeny from Jordan. The genetic causes were identified in six families: two families' due to disease-causing deletions and/or duplication and four families due to missense mutations (3 recessive homozygous, and 1 dominant heterozygous) in genes that are previously implicated with ID. Also, candidate recessive ID genes in other four families were highlighted for the first time and further investigations will be needed to assess the role of these genes in recessive subtypes of ID. While ID in consanguineous families suggest single gene mutations as a genetic cause, the obtained results show that other genetic aberration, such as autosomal dominant copy number variants should also be considered in consanguineous families. Together with the lack of genetic research being performed in Jordan, the high prevalence of consanguinity, along with the high success rate in genetically diagnosing ID patients from this small pilot cohort, the need for larger scale genetic discovery efforts in Jordan and surrounding areas is strongly needed.

Acknowledgments

This study was supported by the Institute for Genomic Medicine (IGM) at Columbia University Medical Center (CUMC) in New York. The binational Fulbright Commission in Jordan supported Tawfiq Froukh's visit to CUMC during this study. The author is grateful to: (1) the affected individuals and their families for their participation in the study; (2) Dr. David Goldstein, Dr. Erin Heinzen and Xiaolin Zhu from IGM/CUMC and Dr. med Vandana Shashi from Duke University School of Medicine in Durham for their continuous scientific advises through this study; and (3) Dr. Marwan Abu Halawah from Philadelphia University-Jordan for his help recruiting family 1.

The author acknowledges the following individuals or groups for the contributions of control samples: K. Welsh-Bomer; C. Hulette; J. Burke; D. Valle; J Hoover-Fong; N. Sobriera; D. Daskalakis; A. Poduri; D. Marchuk; S. Palmer; R. Buckley; S. Hirose; C. Chen; D. Murdock; M. Hauser; A. Holden; E. Behr; M. Winn; R. Gbadegesin; C. Moylan; A.M. Diehl; M. Abdelmalek; the Duke University Health System Nonalcoholic Fatty Liver Disease Research Database and Specimen Repository; S. Schuman; E. Nading; M. Carrington; C. Woods; Croasdaile Village; K. Schmader; S. McDonald; M. Yanamadala; H. White; G. Nestadt; J. Samuels; Y. Wang; T. Young; K. Whisenhunt; J. Milner; J. McEvoy; A. Need; E. Pras; D. Lancet; Z. Farfel; M. Harms; T. Miller; A. Pestronk; R. Bedlack; S. Delaney; E. Cirulli; K, Grace; R. Brown; N. Shneider; S. Gibson; J. Ravits; A. Gitler; J. Glass; F. Baas; S. Appel; E. Simpson; G. Rouleau; S. Halton, L. Lay; the ALS Sequencing Consortium; members of the Center for HIV/AIDS Vaccine Immunology; the National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI); National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery; and the Epi4K Consortium and Epilepsy Phenome/Genome Project. The collection of control samples and data was funded in part by: Biogen, Inc.; Gilead Sciences, Inc.; Bryan ADRC NIA P30AG028377; B57 SAIC-Fredrick Inc. M11-074; National Institute of Neurological Disorders and Stroke (U01NS077303; U01NS053998); National Institute of Mental Health (RC2MH089915, K01MH098126, R01MH097971); Center for HIV/AIDS Vaccine Immunology (CHAVI) (U19-AI067854); National Institute of Allergy and Infectious Diseases

Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (UM1-AI100645); National Institute of Allergy and Infectious Diseases (1R56AI098588-01A1; Division of Intramural Research); the Duke Chancellor's Discovery Program Research Fund 2014; Bill and Melinda Gates Foundation; the Ellison Medical Foundation New Scholar award AG-NS-0441-08; and the Murdock Study Community Registry and Biorepository.

Conflict of Interest

The authors declare no conflict of interest.

References

Abou Jamra, R., Philippe, O., Raas-Rothschild, A., Eck, S.H., Graf, E., Buchert, R., Borck, G., Ekici, A., Brockschmidt, F.F., Nothen, M.M., Munnich, A., Strom, T.M., Reis, A. & Colleaux, L. (2011a) Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature. *Am. J. Hum. Genet.*, **88**, 788-795.

Abou Jamra, R., Wohlfart, S., Zweier, M., Uebe, S., Priebe, L., Ekici, A., Giesebrecht, S., Abboud, A., Alkhateeb, M.A., Fakher, M., Hamdan, S., Ismael, A., Muhammad, S., Nöthen, M.M., Schumaecher, J., et al. (2011b) Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity. *Eur. J. Hum. Genet.*, **19**, 1161-1166.

Akizu, N., Shembesh, N.M., Ben-Omran, T., Bastaki, L., Al-Tawari, A., Zaki, M.S., Koul, R., Spencer, E., Rosti, R.O., Scott, E., Nickerson, E., Gabriel, S., da Gente, G., Li, J., Deardorff, M.A., et al. (2013) Whole-exome sequencing identifies mutated c12orf57 in recessive corpus callosum hypoplasia. *Am. J. Hum. Genet.*, **92**, 392-400.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A. & Shendure, J. (2011) Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.*, **12**, 745-755.

Chinen, Y., Tohma, T., Izumikawa, Y., Uehara, H. & Ohta, T. (2005) Sanfilippo type B syndrome: five patients with an R565P homozygous mutation in the α-N-acetylglucosaminidase gene from the Okinawa islands in Japan. *J. Hum. Genet.*, **50**, 357-359.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. & Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* (*Austin*), **6**, 80-92.

Curtis, D., Vine, A.E. & Knight, J. (2008) Study of regions of extended homozygosity provides a powerful method to explore haplotype structure of human populations. *Ann. Hum. Genet.*, **72**, 261-278.

de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., del Rosario, M., Hoischen, A., Scheffer, H., de Vries, B.B., Brunner, H.G., et al. (2012) Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.*, **367**, 1921-1929.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.*, **43**, 491-498.

Diskin, S.J., Li, M., Hou, C., Yang, S., Glessner, J., Hakonarson, H., Bucan, M., Maris, J.M. & Wang, K. (2008) Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res.*, **36**, e126-e126.

Gibson, J., Morton, N.E. & Collins, A. (2006) Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.*, **15**, 789-795.

Gussow, A.B., Petrovski, S., Wang, Q., Allen, A.S. & Goldstein, D.B. (2016) The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. *Genome Biol.*, **17**, 9.

Hamamy, H.A., Masri, A.T., Al-Hadidy, A.M. & Ajlouni, K.M. (2007) Consanguinity and genetic disorders. Profile from Jordan. *Saudi Med. J.*, **28**,1015-1017.

Han, C., Alkhater, R., Froukh, T., Minassian, A.G., Galati, M., Liu, R.H., Fotouhi, M., Sommerfeld, J., Alfrook, A.J., Marshall, C., Walker, S., Bauer, P., Scherer, S.W., Riess, O., Buchert, R., et al. (2016) Epileptic Encephalopathy Caused by Mutations in the Guanine Nucleotide Exchange Factor DENND5A. *Am. J. Hum. Genet.*, **99**, 1359-1367.

Johansen, A., Rosti, R.O., Musaev, D., Sticca, E., Harripaul, R., Zaki, M., Caglayan, A.O., Azam, M., Sultan, T., Froukh, T., Reis, A., Popp, B., Ahmed, I., John, P., Ayub, M., et al. (2016) Mutations in MBOAT7, Encoding Lysophosphatidylinositol Acyltransferase I, Lead to Intellectual Disability Accompanied by Epilepsy and Autistic Features. *Am. J. Hum. Genet.*, **99**, 912-916.

Kaufman, L., Ayub, M. & Vincent, J.B. (2010) The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodev. Disord.*, **2**, 182-209.

Khoury, S.A. & Massad, D. (1992) Consanguineous marriage in Jordan. *Am. J. Med. Genet.*, **43**, 769-775.

Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., Keerthikumar, S., Oortveld, M.A., Kleefstra, T., Kramer, J.M., Webber, C., Huynen, M.A. & Schenck, A. (2016) Systematic phenomics analysis deconvolutes genes mutated in Intellectual disability into biologically coherent modules. *Am. J. Hum. Genet.*, **98**, 149-164.

Leonard, H. & Wen, X. (2002) The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment. Retard. Dev. Disabil. Res. Rev.*, **8**, 117-134.

Li, H. & Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589-595.

Mefford, H.C., Bastshaw, M.L. & Hoffman, E.P. (2012) Genomics, intellectual disability, and autism. *N. Engl. J. Med.*, **366**, 733-743.

Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., Hosseini, M., Behjati, F., Haas, S., Jamali, P., Zecha, A., Mohseni, M., Puttmann, L., Vahid, L.N., Jensen, C., et al. (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*, **478**, 57-63.

Nalls, M.A., Simon-Sanchez, J., Gibbs, J.R., Paisan-Ruiz, C., Bras, J.T., Tanaka, T., Matarin, M., Scholz, S., Weitz, C., Harris, T.B., Ferrucci, L., Hardy, J. & Singleton, A.B. (2009) Measures of autozygosity in decline: globalization, urbanization, and its implications for medical genetics. *PLoS Genet.*, **5**, e1000415.

Perrault, I., Rozet, J.M., Calvas, P., Gerber, S., Camuzat, A., Dollfus, H., Chatelin, S., Souied, E., Ghazi, I., Leowski, C., Bonnemaison, M., Le Paslier, D.L., Frezal, J., Dufier, J.L., Pittler, S., et al. (1996) Retinal-specific guanylate cyclase gene mutations in Leber's congenital amaurosis. *Nat. Genet.*, **14**, 461-464.

Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S. & Goldstein, D.B. (2013) Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.*, **9**, e1003709.

Poirier, K., Saillour, Y., Bahi-Buisson, N., Jaglin, X.H., Fallet-Bianco, C., Nabbout, R., Castelnau-Ptakhine, L., Roubertie, A., Attie-Bitach, T., Desguerre, I., Geneviève, D., Barnerias, C., Keren, B., Lebrun, N., Boddaert, N., et al. (2010) Mutations in the neuronal β-tubulin subunit TUBB3 result in

malformation of cortical development and neuronal migration defects. *Hum. Mol. Genet.*, **19**, 4462-4473.

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., Dufke, A., Cremer, K., Hempel, M., Horn, D., Hoyer, J., et al. (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet.*, **380**, 1674-1682.

Reuter, M.S., Tawamie, H., Buchert, R., Hosny, G.O., Froukh, T., Thiel, C., Uebe, S., Ekici, A.B., Krumbiegel, M., Zweier, C., Hoyer, J., Eberlein, K., Bauer, J., Scheller, U., Strom, T.M., et al. (2017) Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. *JAMA Psychiatry*, **74**, 293-299.

Schuurs-Hoeijmakers, J.H., Hehir-Kwa, J.Y., Pfundt, R., van Bon, B.W., de Leeuw, N., Kleefstra, T., Willemsen, M.A., van Kessel, A.G., Brunner, H.G., Veltman, J.A., van Bokhoven, H., de Brouwer, A.P. & de Vries, B.B. (2011) Homozygosity mapping in outbred families with mental retardation. *Eur. J. Hum. Genet.*, **19**, 597-601.

Smith, A.C.M., McGavran, L., Robinson, J., Waldstein, G., Macfarlane, J., Zonona, J., Reiss, J., Lahr, M., Allen, L. & Magenis, E. (1986) Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am. J. Med. Genet.*, **24**, 393-414.

Wang, K., Chen, Z., Tadesse, M.G., Glessner, J., Grant, S.F., Hakonarson, H., Bucan, M. & Li, M. (2008) Modeling genetic inheritance of copy number variations. *Nucleic Acids Res.*, **36**, e138.

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. & Bucan, M. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*, **17**, 1665-1674.

Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M.R., Leduc, M.S., Kirby, A., et al. (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.*, **369**, 1502-1511.