Received: 6 August 2018 | Revised: 5 October 2018 | Accepted: 8 October 2018

DOI: 10.1111/cge.13463

WILEY CLINICAL GENETICS

ORIGINAL ARTICLE

Effect of inbreeding on intellectual disability revisited by trio sequencing

Kimia Kahrizi[1] | Hao Hu[2*] | Masoumeh Hosseini[1] | Vera M. Kalscheuer[2] | Zohreh Fattahi[1] | Maryam Beheshtian[1] | Vanessa Suckow[2] | Marzieh Mohseni[1] | Bettina Lipkowitz[2] | Sepideh Mehvari[1] | Zohreh Mehrjoo[1] | Tara Akhtarkhavari[1] | Zhila Ghaderi[1] | Maryam Rahimi[1] | Sanaz Arzhangi[1] | Payman Jamali[3] | Milad Falahat Chian[1] | Pooneh Nikuei[4] | Farahnaz Sabbagh Kermani[5] | Farnaz Sadeghinia[1] | Roshanak Jazayeri[6] | S. Hassan Tonekaboni[7] | Atefeh Khoshaeen[8] | Haleh Habibi[9] | Fatemeh Pourfatemi[10] | Faezeh Mojahedi[11] | Mohammad-Reza Khodaie-Ardakani[12] | Reza Najafipour[13] | Thomas F. Wienker[2] | Hossein Najmabadi[1,14] | Hans-Hilger Ropers[2,15]

[1]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[2]Max Planck Institute for Molecular Genetics, Berlin, Germany

[3]Shahrood Genetic Counseling Center, Welfare Office, Semnan, Iran

[4]Molecular Medicine Research Center, Hormozgan Health Institute, Hormozgan University of Medical Sciences, Bandar Abbas, Iran

[5]Clinical Research Unit, Afzalipour Hospital, University of Medical Sciences, Kerman, Iran

[6]Department of Biochemistry, Genetic and Nutrition, Faculty of Medicine, Alborz University of Medical Sciences, Karaj, Iran

[7]Pediatric Neurology Research Center, Mofid Children's Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[8]Genetic Consultant, Sari, Iran

[9]Hamedan University of Medical Science, Hamedan, Iran

[10]Sari Genetic Counseling Center, Welfare Office, Sari, Iran

[11]Mashhad Medical Genetic Counseling Center, Mashhad, Iran

[12]Razi Hospital, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[13]Cellular and Molecular Research Centre, Genetic Department, Qazvin University of Medical Sciences, Qazvin, Iran

[14]Kariminejad - Najmabadi Pathology and Genetics Center, Tehran, Islamic Republic of Iran

[15]Institute for Human Genetics, University Medicine Mainz, Germany

Correspondence
Hans-Hilger Ropers, Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, D-14195 Berlin, Germany.
Email: ropers@molgen.mpg.de

[*]**Present address**
Hao Hu, Guangzhou Institute of Pediatrics, Guangzhou Women and Children's Medical Center, 9 Jinsui Road, Guangzhou, Guangdong, 510623, P.R.China.

Funding information
Iran National Science Foundation, Grant/Award Numbers: grant no. 92038458, 950022,

In outbred Western populations, most individuals with intellectual disability (ID) are sporadic cases, dominant *de novo* mutations (DNM) are frequent, and autosomal recessive ID (ARID) is very rare. Because of the high rate of parental consanguinity, which raises the risk for ARID and other recessive disorders, the prevalence of ID is significantly higher in near- and middle-east countries. Indeed, homozygosity mapping and sequencing in consanguineous families have already identified a plethora of ARID genes, but because of the design of these studies, DNMs could not be systematically assessed, and the proportion of cases that are potentially preventable by avoiding consanguineous marriages or through carrier testing is hitherto unknown. This prompted us to perform whole-exome sequencing in 100 sporadic ID patients from Iran and their healthy consanguineous parents. In 61 patients, we identified apparently causative changes

Kimia Kahrizi and Hao Hu, as well as Hossein Najmabadi and Hans-Hilger Ropers, contributed equally to this study.

92035782; The National Institute for Medical Research Development, Grant/Award Number: 957060; the European Union through FP7 project GENCODYS, Grant/Award Number: 241995; Max Planck Innovation Funds

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1111/cge.13463/

in known ID genes. Of these, 44 were homozygous recessive and 17 dominant DNMs. Assuming that the DNM rate is stable, these results suggest that parental consanguinity raises the ID risk about 3.6-fold, and about 4.1 to 4.25-fold for children of first-cousin unions. These results do not rhyme with recent opinions that consanguinity-related health risks are generally small and have been "overstated" in the past.

KEYWORDS
impact of inherited and de novo mutations, intellectual disability risks, parental consanguinity, parent-patient trios, whole-exome sequencing

1 | INTRODUCTION

Parental consanguinity (PC) is known to raise the risk of having children with recessive disorders, and it is associated with increased prenatal and pre-reproductive mortality. In the offspring of first-cousin matings, the prevalence of major congenital malformations (CM) is 2% to 2.5% higher than in children of unrelated parents; in Western countries, this more than doubles the CM risk ([1–5] and refs therein). Moreover, 3- to 5-fold elevated intellectual disability (ID) risks have been reported for children whose parents are first cousins,[6–10] but these estimates are based on few and mostly small studies with different designs, rendering direct comparisons difficult.

PC and autosomal recessive forms of ID (ARID) are rare in most parts of Europe and the United States, where the vast majority of ID patients are sporadic cases and dominant *de novo* mutations (DNMs) are common.[11–13] In contrast, PC is common in many developing countries, from Morocco through the near- and middle-east to parts of India (see Ref. 14), with current consanguinity rates being highest in Pakistan and Sudan. Recessive disorders are also a major health care problem in many other countries of the so-called "consanguinity belt," including Iran and, for example, Qatar, where a 30-year health plan aims to reduce the frequency of consanguineous matings.[15]

As shown a decade ago,[16] ARID is extremely heterogeneous, and the number of causative genes is likely to run into thousands.[17] This has been confirmed by large-scale, high-throughput sequencing in consanguineous families with two or more intellectually disabled children (see Ref. 18 and refs therein): between the first and the most recent study of this kind,[18,19] the percentage of multiplex families with mutations in known ARID genes has only risen from one-third to about one half[18] while the number of known ARID genes has increased from about a dozen to about 800 today.[18,20,21]

In a recent study of the British Deciphering Developmental Defects (DDD) project,[22] whole-exome sequencing (WES) identified bi-allelic (recessive) mutations in 3.6% of the patients with developmental disorders (DD) and European ancestry, compared to 49.9% DNMS in known DD genes. To date, no comparable molecular data have been published for countries where PC is common, because most sequencing studies performed in Iran, Turkey, Saudi Arabia or Pakistan focused on recessive gene defects and were not designed to capture DNM (eg,[19–21]), did not quantify consanguinity or autozygosity, or were simply too small[23].

Therefore, and in view of persisting disputes about the size of health care risks conferred by PC and the appropriateness and efficiency of efforts to discourage consanguineous matings (eg, see Refs. 24 and 25), we present here the results of molecular investigations designed to infer the proportion of DNM and recessive inherited mutations causing ID in the offspring of healthy consanguineous parents. These data provide a backbone for estimating preventable ID risks, in countries where PC is common and world-wide.

2 | MATERIAL AND METHODS

2.1 | Patients

The study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation, Tehran, Iran. Consent for participation of their child in this study was obtained from at least one of the parents and if possible, from the patients themselves.

Families with a single intellectually disabled child and healthy consanguineous parents were recruited from different provinces of Iran, who had been pre-screened to exclude chromosomal rearrangements as well as fragile X syndrome. In total, 100 unrelated ID patients and their parents were included in this study. All underwent a comprehensive clinical examination by experienced clinical geneticists. To determine the intelligence quotient (IQ) of children and their parents, we used Wechsler Intelligence Scales for Children (WISC) and adults (WAIS). Clinical findings in affected children are provided in Table 1 and Table S1.

2.2 | Whole-exome sequencing and array-comparative genomic hybridization

We extracted genomic DNA from white blood cells of patients and parents, enriched exomic sequences with the Agilent SureSelectXT Human All Exon V5 enrichment kit, performed 100 bp paired-end sequencing on an Illumina HiSeq sequencer and employed the MERAP pipeline for sequence analysis, as previously described.[18,26]

Previous studies[16,19] had shown that in consanguineous ARID families, microdeletions or duplications are very rare, and that the analysis of WES data with our previously described MERAP pipeline[26] identifies most clinically relevant homozygous copy number variants (CNVs). Therefore, array-based comparative genomic hybridization

(array CGH) was only performed in individuals where WES had failed to identify possible disease-causing variants. For these studies, we employed the CYTOCHIP ISCA 8X60K whole-genome oligo array version 2. This array includes 60 000 spots with 51 Kb mean backbone resolution and close to 500 targeted disease regions. The samples were hybridized twice against reference samples, and the BlueFuse Multi software was used to analyze the data.

2.3 | Criteria for selecting ID-causing mutations

Loss-of-function mutations in known ID genes were accepted as disease-causing if the phenotype of the affected individual matched previous descriptions; missense mutations were considered as disease-causing if their ACMG score, determined by InterVar/wInterVar analysis[27] was pathogenic or probably pathogenic, using annotations generated by ANNOVAR[28]; if three out of four pathogenicity scores (SIFT, PolyPhen2, MutationTaster and CADD) as well as Logit scores provided by MERAP were high, and/or if the phenotype matched and was specific enough to establish the molecular diagnosis. For all variants selected in families with healthy children, we also performed cosegregation studies. Possible ARID mutations were only considered if embedded in a sufficiently large run of homozygous markers, essentially using the ROH definitions in[29] (for details, see Supporting Information). Previous analyses had revealed that in general, the density of informative variants retrieved by WES is high enough to reliably identify ROHs in the offspring of consanguineous parents (eg, see[18]). A comprehensive overview of the criteria for selecting mutations causing ID is provided in Table S1.

3 | RESULTS

3.1 | Mutations involving known ID genes

All mutations in previously described ID genes were only seen in a single patient; for two genes, *SUCLA2* and *ATP8A2*, allelic mutations were observed in two separate ID families. In another family (M8800167), the complex ID phenotype may result from loss of function (LOF) mutations in two genes, a homozygous frameshift mutation inactivating *HADH*, a known ID gene, as well as a heterozygous *de novo*, stop-gain mutation in *DRD2*, which may explain the psychotic signs observed in this patient. *De novo DRD2* mutations have been reported in patients with behavioral abnormalities, but not with ID (for refs see Table S1).

Of note, we found a potentially actionable mutation inactivating *TRPM6* in a 9-year-old girl with seizures, muscle spasms and autistic behavior (family M9300155). *TRPM6* plays an essential role in epithelial magnesium transport and absorption, and pathogenic *TRPM6* mutations may cause infantile epileptic encephalopathy. Magnesium supplementation has been reported to restore normal brain function if started early in life.[30]

For many of the genes listed in Table S1 (see families M9300163, M9300035, M9300014, M8800189 and M8900019), mutations in patients with ID have only been reported very recently, including *VPS13C*[31]; *BOD1*,[32] *TTI1*,[33] *NEURL4*[18]; *ATP13A1*,[34] a paralog of *ATP13A2*, which is associated with autosomal recessive early-onset parkinsonism, ceroid lipofuscinosis and ID,[35] as well as *MED13*.[36] Thus, our study confirms their identity as ID genes. Table S1 also lists three apparently pathogenic CNVs that were detected by array-CGH after WES had failed to identify disease-causing variants in these patients. Details about these CNVs and the relevant genes are provided in the Supporting Information.

Taken together, our analyses suggest that we have identified the molecular cause of ID in 61 (61%) of these trios (see Tables 1 and S1). In 26 trios, we observed rare disruptive or damaging variants involving brain-expressed genes that had not been associated with ID before, including four genes that have been implicated in other disorders. Table 1 summarizes our findings in families with ADID, ARID and XLID. Compound heterozygous mutations were not observed (for further details, see Table S1).

3.2 | Novel candidate genes for ID

Several of the selected variants in the remaining families were not listed in the ExAC and gnomAD databases, and for most others, allele frequencies were very low. As recently reported,[37] the frequency of ultra-rare disruptive or damaging mutations in brain-expressed genes correlates with a decrease in years of education, a proxy for low IQ. In most of the unsolved trios, stringent filtering enabled us to reduce the number of possibly pathogenic variants in plausible candidate genes to a single one, but occasionally, two or even three possibly ID-causing variants in novel candidate genes were retained. Thus, numerous but not all of the candidate genes listed in Table S1 may be novel ID genes. Supportive evidence for several particularly promising ID candidate genes is provided in the Supporting Information, and for a number of others, allelic mutations in unrelated ID patients were identified through Gene Matcher-mediated data exchange with other groups.[38] Detailed genotypes and phenotypes of these matching cases will be published separately. When estimating the frequency of dominant DNMs and recessive inherited mutations in our cohort, variants in novel candidate genes were not included to avoid unnecessary biases.

3.3 | Inferring ID risks from ARID and ADID frequencies in consanguineous families

For this analysis, we have compared the frequency of ARID and ADID among the "solved" cases with mutations in established ID genes.

Affected individuals from 44 trio families carried inherited ID-causing mutations, including 41 homozygous mutations in known ARID genes, two hemizygous mutations in known XLID genes, and a homozygous CNV truncating a gene implicated in a syndromic form of ID. Pathogenic DNM were identified in 17 individuals with ID, 14 in known ADID genes and one in a known XLID gene, and 2 of the *de novo* mutations were CNVs. Thus, 44 out of 61 disease-causing mutations in "solved" trios (71.13%) turned out to be inherited, while only 17 (27.86%) were DNM.

In keeping with other studies of consanguineous families, most but not all couples of our cohort are first cousins. While the coefficient of consanguinity for offspring of first-cousin matings is

TABLE 1 Clinical features and mutation spectrum observed in 100 individuals with sporadic ID and consanguineous parents

Family_Number	ID or DD	Microcephaly (<= — 2.0 SD)	Epilepsy	Autistic features	Additional features	Gene_ Symbol [HGNC]	ID_Gene_ Status	Mutation_Status	Mutation consequence	Change_Protein [HGVS]
M9300156	Moderate ID	MA	N	N	Y	ARHGAP32	Known	DNM	stop_gained	NP_001136157.1:p.(Arg724Ter)
M9300052	Moderate ID	−2.0 SD	Y	N	Y	BCL11A	Known	DNM	frameshift_variant	NP_075044.2:p.(Leu61SerfsTer23)
M9300056	Moderate ID	−3.0 SD	N	N	Y	CD163L1	Unclear	DNM	missense_variant	NP_001284579.1:p.(Gly1379Arg)
M8900244	Mild ID	−6.5 SD	N	N	Y	DICER1		DNM	missense_variant	NP_803187.1:p.(Thr647Ile)
M9300115	DD	−8.4 SD	Y	N	Y	DLX6	Novel	DNM	inframe_deletion	NP_005213.3:p.(Gln42_Gln44del)
M9300158	Severe ID	N	N	N	Y	DNM1L	Known	DNM	missense_variant	NP_001265393.1:p.(Leu649Pro)
M8800167	Moderate ID	N	Y	N	Y	DRD2		DNM	stop_gained	NP_000786.1:p.(Arg275Ter)
M8800055	Moderate ID	−2.0 SD	N	N	Y	FBXO11	Known	DNM	splice_acceptor_variant	—
M9300104	Moderate ID	N	Y	N	Y	GRIN2B	Known	DNM	missense_variant	NP_000825.2:p.(Arg682His)
M9300159	Mild ID	MA	N	Y	Y	GRM7	Known	DNM	stop_gained	NP_870989.1:p.(Arg861Ter)
M9300105	Moderate ID	N	N	N	Y	HTR2A	Novel	DNM	missense_variant	NP_000612.1:p.(Leu136Arg)
M9300085	Severe ID	−3.5 SD	N	N	Y	KAT6A	Known	DNM	stop_gained	NP_001092882.1:p.(Glu1234Ter)
M9300008	Mild ID	−5.0 SD	N	N	Y	KIF11	Known	DNM	stop_gained	NP_004514.2:p.(Arg47Ter)
M9300026	Moderate ID	N	Y	N	Y	KMT2A	Known	DNM	missense_variant	NP_001184033.1:p.(Arg1154Gln)
M9300160	Moderate ID	−4.0 SD	N	N	Y	MED13	Known	DNM	frameshift_variant	NP_005112.2:p.(Ala443GlufsTer6)
M9300057	Mild ID	N	Y	N	Y	NFIB	Novel	DNM	missense_variant	XP_005251524.1:p.(Asp119His)
M9300006	Moderate ID	−3.5 SD	Y	N	Y	OCM2	Novel	DNM	missense_variant	NP_006179.2:p.(Asp60Gly)
M9300036	Mild ID	N	N	N	Y	PIWIL1	Novel	DNM	missense_variant	NP_004755.2:p.(Arg589Gln)
M9300090	Moderate ID	−5.0 SD	N	N	Y	RCAN1	Novel	DNM	missense_variant	XP_005260986.1:p.(Gly147Arg)
M9300031	Moderate ID	N	N	N	Y	TCF20	Known	DNM	frameshift_variant	NP_005641.1:p.(Gln1106ArgfsTer3)
M011	Mild ID	−2.9 SD	N	N	Y	TRIO	Known	DNM	frameshift_variant	NP_009049.2:p.(Val1698LeufsTer61)
M9300038	Mild ID	N	N	N	Y	TRIP12	Known	DNM	missense_variant	NP_001271143.1:p.(Pro466Ala)
M9300064	Severe ID	−2.5 SD	N	N	Y	XRRA1	Novel	DNM	frameshift_variant	XP_005273822.1:p.(Thr758SerfsTer16)
M9300080	DD	−3.0 SD	Y	N	Y	ZEB2	Known	DNM	stop_gained	NP_055610.1:p.(Gln368Ter)
M9300016	Moderate ID	Y;-3,3SD	N	N	Y		Known	DNM	CNV	—
M9300010	Profound ID	−6.0 SD	Y	N	Y	ADAM28	Novel	INH	missense_variant	NP_055080.2:p.(Gly480Ser)
M9300018	Moderate ID	N	Y	N	Y	ADGRG1	Known	INH	missense_variant	NP_001139243.1:p.(Arg565Trp)
M9300051	Severe ID	−3.5 SD	N	N	Y	AHSG	Known	INH	stop_gained	NP_001613.2:p.(Lys2Ter)
M9300065	Severe ID	N	N	N	Y	AIMP1	Known	INH	missense_variant, splice_region_variant	NP_001135888.1:p.(Gly282Arg)
M9000002	Moderate ID	N	Y	N	Y	ALDH5A1	Known	INH	stop_gained, frameshift_variant	NP_733936.1:p.(Tyr47Ter)
M9300122	DD	−9.0 SD	Y	N	Y	ALPK1	Novel	INH	missense_variant	XP_005263302.1:p.(Leu128Met)
M340	Moderate ID	−5.6 SD	Y	N	Y	AMHR2	—	INH	stop_gained	NP_065434.1:p.(Arg22Ter)
M8800054	Severe ID	−4.0 SD	N	N	Y	AP4M1	Known	INH	missense_variant	XP_005250746.1:p.(His340Pro)
M8900020	Profound ID	−2.8 SD	N	N	Y	ASAH1	Known	INH	missense_variant, splice_region_variant	NP_004306.3:p.(Trp185Arg)

TABLE 1 (Continued)

Family_Number	ID or DD	Microcephaly (<= — 2.0 SD)	Epilepsy	Autistic features	Additional features	Gene_ Symbol [HGNC]	ID_Gene_ Status	Mutation_Status	Mutation consequence	Change_Protein [HGVS]
M9300103	DD	−3.0 SD	Y	N	Y	ASNS	Known	INH	missense_variant	NP_899199.2:p.(Ile253Ser)
M8700030	Moderate ID	−11 SD	N	N	N	ASPM	Known	INH	frameshift_variant	NP_060606.3:p.(Lys633IlefsTer8)
M8900019	Severe ID	N	N	N	Y	ATP13A1	Known	INH	missense_variant	XP_005260049.1:p.(Gly547Cys)
M9300050	Moderate ID	−3.0 SD	Y	N	Y	ATP8A2	Known	INH	splice_donor_variant	—
M9300178	DD	N	N	N	Y	ATP8A2	Known	INH	stop_gained	NP_057613.4:p.(Trp1049Ter)
M9300035	Moderate ID	N	N	N	Y	BOD1	Known	INH	frameshift_variant	NP_612378.1:p.(Gly63AlafsTer23)
M9300154	Mild ID	N	Y	N	Y	CDK10	Known	INH	splice_acceptor_variant	—
M8700062	Moderate ID	−7.5 SD	N	N	Y	CDK5RAP2	Known	INH	frameshift_variant	NP_060719.4:p.(Ile1144GlnfsTer13)
M8700060	Mild ID	−7.7 SD	N	N	N	CENPF	Known	INH	stop_gained	NP_057427.3:p.(Gln541Ter)
M9300021	Mild ID	−5.0 SD	N	N	N	CENPJ	Known	INH	missense_variant	NP_060921.3:p.(Lys150Glu)
M9300063	Moderate ID	MA	N	Y	Y	CFAP44	—	INH	splice_acceptor_variant	—
M9300017	Severe ID	−4.7 SD	N	N	Y	CHMP1A	Known	INH	intron_variant	—
M9300066	Mild ID	N	N	N	N	CKAP5	Novel	INH	missense_variant	NP_001008938.1:p.(Ala891Thr)
M9300030	Mild ID	−2.9 SD	N	N	Y	DDX50	Novel	INH	inframe_deletion	NP_076950.1:p.(Ser690_Gly698delinsCys)
M9300121	Severe ID	N	Y	N	Y	DENND2A	Novel	INH	missense_variant	NP_056504.3:p.(Arg930Trp)
M9300164	Moderate ID	N	Y	N	Y	DNAH3	Novel	INH	missense_variant	NP_060009.1:p.(Ser3309Leu)
M9300121	Severe ID	N	Y	N	Y	DYRK4	Novel	INH	missense_variant	NP_003836.1:p.(Phe304Leu)
M9300062	Severe ID	N	N	N	Y	EXOSC2	Known	INH	splice_acceptor_variant	—
M9300029	Severe ID	−6.0 SD	Y	N	Y	FUCA1	Known	INH	frameshift_variant	NP_000138.2:p.(Arg354GlyfsTer9)
M9300067	Moderate ID	N	Y	N	Y	FZD3	Novel	INH	missense_variant	NP_059108.1:p.(Phe231Cys)
M9300102	Moderate ID	N	Y	N	Y	GLRA1	Known	INH	CNV	—
M9300024	Mild ID	−3.0 SD	N	N	Y	GMPPA	Known	INH	stop_gained	XP_005246542.1:p.(Cys181Ter)
M8800167	Moderate ID	N	Y	N	Y	HADH	Known	INH	frameshift_variant	NP_001171634.2:p.(Lys206SerfsTer14)
M9300162	Moderate ID	N	N	N	Y	HIST1H1C	Novel	INH	missense_variant	NP_005310.1:p.(Thr45Ile)
M9300116	Severe ID	−4.5 SD	N	N	Y	KIF7	Known	INH	missense_variant	NP_940927.2:p.(Asn87Asp)
M9300119	Severe ID	−3.5 SD	N	N	Y	LDB3	—	INH	missense_variant	NP_001165081.1:p.(Ile563Val)
M9300126	DD	−4.5 SD	Y	N	Y	LPIN3	Novel	INH	missense_variant	NP_001288789.1:p.(Arg59Trp)
M8900015	Profound ID	−3.3 SD	N	Y	Y	LYST	Known	INH	missense_variant	NP_001288294.1:p.(Asp2467Gly)
M9300053	Severe ID	N	N	N	Y	MYOCD	Known	INH	missense_variant	NP_001139784.1:p.(Leu629Ile)
M8800189	Moderate ID	−3.5 SD	N	N	Y	NEURL4	Known	INH	stop_gained	NP_115818.2:p.(Gln1096Ter)
M9300129	Moderate ID	N	N	N	Y	P4HA3	Novel	INH	missense_variant	NP_878907.1:p.(Asp505His)
M9300025	Moderate ID	N	N	N	Y	PEPD	Known	INH	missense_variant	NP_000276.2:p.(His255Pro)
M8800093	DD	−2.0 SD	N	N	Y	PLA2G6	Known	INH	missense_variant	XP_005261823.1:p.(Arg329Cys)
M8900086	Profound ID	−6.0 SD	N	N	Y	POLRMT	Novel	INH	missense_variant	XP_005259637.1:p.(Phe692Leu)
M8800040	Severe ID	−4.2 SD	N	N	Y	PTRH2	Known	INH	frameshift_variant	NP_001015509.1:p.(Gln86LeufsTer40)

TABLE 1 (Continued)

Family_Number	ID or DD	Microcephaly (<= — 2.0 SD)	Epilepsy	Autistic features	Additional features	Gene_ Symbol [HGNC]	ID_Gene_ Status	Mutation_Status	Mutation consequence	Change_Protein [HGVS]
M9300049	Mild ID	N	N	N	Y	RALGAPA1	Novel	INH	missense_variant	XP_005267548.1:p.(Arg2122His)
M9300060	Severe ID	−7.5 SD	N	N	Y	RNASEH2C	Known	INH	missense_variant	NP_115569.2:p.(Pro151Ser)
M9300122	DD	−9.0 SD	Y	N	Y	RPLP1	Novel	INH	inframe_deletion, splice_region_variant	NP_000994.1:p.(Glu91del)
M9300053	Severe ID	N	N	N	Y	SCG2	Novel	INH	missense_variant	NP_003460.2:p.((Glu382Lys)
M9300091	DD	−7.5 SD	N	N	N	SCGN	Novel	INH	missense_variant	NP_008929.2:p.(Leu268Ser)
M9300033	Mild ID	N	N	N	Y	SIK1	Known	INH	missense_variant	NP_775490.2:p.(Pro420Leu)
M9300011	Moderate ID	N	Y	Y	Y	SIX4	Novel	INH	missense_variant	NP_059116.3:p.(Ala237Thr)
M9300032	Moderate ID	N	N	N	Y	SPTBN2	Known	INH	frameshift_variant	XP_005274249.1:p.(His407AlafsTer4)
M9300095	Moderate ID	N	Y	N	Y	SRD5A3	Known	INH	stop_gained	NP_078868.1:p.(Trp19Ter)
M9300115	DD	−8.4 SD	Y	N	Y	STXBP5	Novel	INH	missense_variant	NP_001121187.1:p.(Asp489Gly)
M9300088	DD	−3.3 SD	N	N	Y	SUCLA2	Known	INH	missense_variant	NP_003841.1:p.(Arg284Cys)
M9300093	Severe ID	−2.3 SD	N	N	Y	SUCLA2	Known	INH	missense_variant	NP_003841.1:p.(Thr145Pro)
M9300076	Moderate ID	N	N	N	Y	TMEM67	Known	INH	missense_variant	NP_714915.3:p.(Asn242Ser)
M9300155	Moderate ID	N	Y	Y	Y	TRPM6	Known	INH	stop_gained	NP_060132.3:p.(Ser1034Ter)
M9300047	Moderate ID	−2.0 SD	Y	N	N	TRRAP	Known	INH	missense_variant	XP_005250674.1:p.(Thr60Ile)
M9300089	Moderate ID	−3.5 SD	N	N	Y	TSEN54	Known	INH	missense_variant, splice_region_variant	XP_005257286.1:p.(Gly124Val)
M9300014	Mild ID	−3.3 SD	Y	N	Y	TTI1	Known	INH	missense_variant, splice_region_variant	NP_055472.1:p.(Leu767Ser)
M9300122	DD	−9.0 SD	Y	N	Y	TUBGCP2	Novel	INH	missense_variant	NP_001243546.1:p.(Ala643Pro)
M9300087	Mild ID	N	N	N	Y	UBE3B	Known	INH	missense_variant	NP_569733.2:p.(Asp14Asn)
M9300163	Mild ID	N	N	N	Y	VPS13C	Known	INH	stop_gained	NP_065872.1:p.(Cys2627Ter)
M8800142	DD	−3.0 SD	N	N	Y	YTHDF1	Novel	INH	missense_variant	XP_005260267.1:p.(Gly13Arg)
M9300063	Moderate ID	MA	N	Y	Y	ZNF804A	Novel	INH	missense_variant	NP_919226.1:p.(Ala55Thr)
X-linked ID										
M9300039	Mild ID	N	N	N	N	DDX3X	Known	DNM	missense_variant	NP_001347.3:p.(Val345Glu)
M9300166	Mild ID	−2.5 SD	N	N	Y	HUWE1	Known	DNM	CNV	—
M8800151	Moderate ID	−2.7 SD	N	N	Y	PQBP1	Known	INH	frameshift_variant	NP_005701.1:p.(Arg153SerfsTer41)
M9300037	Mild ID	−2.0 SD	N	Y	Y	TAF1	Known	INH	missense_variant	NP_001273003.1:p.(Pro133Arg)

Abbreviations: CNV, copy number variants; DD, developmental disorders; DNM, de novo mutations; HGNC, Hugo Gene Nomenclature committee; HGVS, Human Genome Variation Society; ID, intellectual disability; INH, inherited; N, no; Y, yes.

F = 0.0625, the arithmetic mean of F for the 100 probands of our cohort was only 0.05256 (see Table S1). Therefore, the 44 cases with inherited forms of ID observed in our cohort would correspond with 44 × 0.0625/0.05256 = 52.32 cases if all parents were first cousins, and the ID risk of their children would be (52.32 + 17)/17 = 4.08 times higher than the risk of children with unrelated parents. Among the 34 "solved" cases with first-cousin parents of our cohort (F = 0.0625, see Table S1), 26 (76.47%) were homozygous for inherited ARID mutations while only 8 (23.53%) carried DNM, corresponding with a 4.25-fold higher ID risk.

4 | DISCUSSION

4.1 | Impact of inbreeding on the incidence of ID

Numerous studies have estimated the impact of inbreeding on the incidence of ID, with varying results (eg,[7,9,39–41]). Most published risk estimates were lower than presented here, but it is not straightforward to compare these diverging results because they were obtained in different ways and different populations. To the best of our knowledge, our study is the first to employ WES data from patients with consanguineous parents for this purpose. However, there are several reasons why the outcome of this approach may be biased, too, as discussed below.

First, for example, we have used pedigree information to infer the F values for the affected individuals. However, in populations with a long tradition of marrying within the family, hidden consanguinity, too distant to be visible in the pedigree, cannot be excluded. Indeed, ROH studies in Iran and other middle-east countries have revealed traces of ancient consanguinity in purportedly unrelated couples.[42] Thus, it is probable that we have underestimated the consanguinity in some or even many of the families studied, and consequently, the impact of the "visible" consanguinity on the incidence of ID may have been overestimated. For the absolute ID risk of children born to consanguineous Iranian couples, this would have no consequences, but for offspring of apparently unrelated Iranian parents, hidden consanguinity would enhance the ID risk. In contrast, the ID risks of consanguineous families should be lower in outbred (eg, Central European) populations where hidden consanguinity is expected to be rare or absent.

Second, as shown by Wahlund[43] and referred to by Overall,[44] the existence of endogamous subpopulations, for which there is ample evidence in Iran and neighboring countries, will lead to an excess of homozygosity. This may reduce the proportion of cases that can be prevented by avoiding consanguineous marriages,[45] but these effects are generally small. In the present study, significant "Wahlund effects" are also unlikely because all homozygous mutations considered as causative for ARID were embedded in (mostly large) ROHs, indicating that the paternally and maternally transmitted alleles were identical by descent and derived from a closely related common ancestor.

Third, not all ID-causing mutations can be detected by Trio WES, such as deep intronic variants interfering with splicing or remote regulatory mutations controlling gene expression, and some apparently benign exonic variants may in fact be pathogenic (eg, exonic splicing regulators, see Refs. 46–48). While missing these mutations may impact on the proportion of dominant and recessive causes of ID, it is probable that their effect will be relatively small.

In consanguineous pedigrees with two or more affected patients, where ARID is the most probable explanation, mutations in known ID genes account for only 50% of the families. This may reflect the large size of the ARID gene pool, for which there is now compelling empirical evidence (eg, see Hu et al.[18]). Similarly, only 50% of sporadic patients from outbred populations carry DNMs in known ADID genes,[22] but given the evidence that many ADID genes are already known, mutations in hitherto unknown genes cannot explain these findings. Instead, in sporadic cases that do not carry DNMs in known ID genes, other disease-causing mechanisms may apply, including di- or oligogenic inheritance, which has been described in patients with CNVs,[49,50] and even non-genetic (eg, epigenetic, environmental or stochastic) factors may play important roles.[51]

Although most ARID genes are still unknown and the etiology of sporadic ID is still largely unexplored, the available evidence suggests that mutations in presently known ARID genes account for about half of the inherited ID risk, and that DNM account for half the ID risk of sporadic cases. Therefore, the proportion of affected individuals with mutations in known ARID and ADID genes in our cohort should be the same as the ratio between all patients with ARID and all individuals with sporadic ID.

In this study, the proportion of cases with recessively inherited mutations and with DNM was 44 to 17, as shown above. Assuming that DNMs are equally frequent in different populations and not influenced by consanguinity, it follows that for offspring of related Iranian parents, the ID risk will be raised (44 + 17)/17 = 3.59-fold.

4.2 | ARID risks in middle east countries and world-wide

While environmental factors, social deprivation and the quality of health care may influence the IQ, inbreeding has emerged as the most important factor in ID, even in deprived parts of India.[41] Our data suggest that in countries like Iran or Qatar, where close to 40% of the children have consanguineous parents, the genetic ID risk should be twice as high as in Central Europe, where consanguinity rates are negligibly low, and in areas of Pakistan or Saudi Arabia with up to 70% consanguineous marriages, the genetic ID risk (and hence, the incidence of ID) should be almost three times as high.

According to our data, the ID risk is even higher for children of first cousins, which is in keeping with recent results of a study conducted in India.[41] Even though in Western populations, the consanguinity-related risks may be slightly lower because of the absence of hidden consanguinity, they are still considerable and certainly higher than ID risks related to advanced parental age.

5 | CONCLUSION AND OUTLOOK

This study has shown that in the offspring of healthy consanguineous parents, ID is mostly due to homozygous recessive mutations, in contrast to the offspring of unrelated parents, where DNM are the

predominant cause of ID. Our data suggest that having first-cousin parents raises the ID risk 4.1- to 4.25-fold, higher than suggested by several recent studies, but in good agreement with the 5-fold rise predicted by Newton Morton 40 years ago.[9] These results should be of interest to related couples and genetic counselors alike, not only in highly consanguineous populations.

Using WES and array-CGH to investigate consanguineous parents and their only affected child also enabled us to detect pathogenic variants affecting known ID genes in 61 out of 100 trios analyzed and to identify several candidate genes for ID that had not been reported before.

Moreover, our results illustrate that ARID is an important global health care problem that is still far from being solved, partly because it is relatively rare in outbred Western populations where most of the research takes place. Given the enormous genetic heterogeneity of ARID, even Genomics England's DDD study is far too small to identify most or all underlying recessive defects. In view of the high diagnostic success rate of trio-WES and -WGS and the recent decision of the English National Health Service to implement (simplex or subsequently trio) WGS in routine genetic health care,[52] the time may be ripe for also introducing these methods as first-line genetic tests in Iran and other highly consanguineous Middle-Eastern populations. As shown here, this would greatly accelerate the identification of the many ARID genes that are hitherto unknown; lay the groundwork for efficient ID prevention through pre-conception carrier testing; and shed more light on the function of the human brain in health and disease.

ACKNOWLEDGEMENTS

We thank all patients and families who participated in this study. Thanks to Gabriele Eder for secretarial assistance. The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. Financial support has been provided by the Iran National Science Foundation (grant nos 92038458, and 950022) to HN, (grant no. 92035782) to KK, and National Institute for Medical Research Development (grant no. 957060) to KK and the European Union through FP7 project GENCODYS, grant no. 241995 (organizer: Hans van Bokhoven) to HHR and HN, and by the Max Planck Innovation Funds (to HHR).

ORCID

Hans-Hilger Ropers https://orcid.org/0000-0002-8552-9813

REFERENCES

1. Bennett RL, Motulsky AG, Bittles A, et al. Genetic counseling and screening of consanguineous couples and their offspring: recommendations of the National Society of genetic counselors. *J Genet Couns*. 2002;11:97-119.
2. Darr A, Small N, Ahmad WIU, Atkin K, Corry P, Modell B. Addressing key issues in the consanguinity-related risk of autosomal recessive disorders in consanguineous communities: lessons from a qualitative study of British Pakistanis. *J Community Genet*. 2016;7:65-79.
3. Jaber L, Halpern GJ, Shohat M. The impact of consanguinity worldwide. *Community Genet*. 1998;1:12-17.
4. Stoll C, Alembik Y, Roth MP, Dott B. Parental consanguinity as a cause for increased incidence of births defects in a study of 238,942 consecutive births. *Ann Genet*. 1999;42:133-139.
5. Zlotogora J. What is the birth defect risk associated with consanguineous marriages? *Am J Med Genet*. 2002;109:70-71.
6. Book JA. Genetical investigations in a north Swedish population: the offspring of first-cousin marriages. *Ann Hum Genet*. 1957;21:191-221.
7. Bundey S, Alam H. A five-year prospective study of the health of children in different ethnic groups, with particular reference to the effect of inbreeding. *Eur J Hum Genet*. 1993;1:206-219.
8. Madhavan T, Narayan J. Consanguinity and mental retardation. *J Ment Defic Res*. 1991;35(Pt 2):133-139.
9. Morton NE. Effect of inbreeding on IQ and mental retardation. *Proc Natl Acad Sci U S A*. 1978;75:3906-3908.
10. Reed EW, Reed SC. *Mental Retardation: A Family Study*. Philadelphia, PA: Saunders; 1965.
11. de Ligt J, Willemsen MH, van Bon BWM, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012;367:1921-1929.
12. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet*. 2012;380:1674-1682.
13. Yuen RKC, Thiruvahindrapuram B, Merico D, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med*. 2015;21:185-191.
14. Bittles AH, Black ML. *Global Patterns & Tables of Consanguinity*. 2015. http://consang.net, accessed: August 04, 2018.
15. Qatar National Health Strategy. (http://www.qu.edu.qa/pharmacy/components/Qatar_National_Health_Strategy.pdf). 2011–2016. accessed August 04, 2018.
16. Najmabadi H, Motazacker MM, Garshasbi M, et al. Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet*. 2007;121:43-48.
17. Ropers HH. Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet*. 2010;11:161-187.
18. Hu H, Kahrizi K, Musante L, et al. Genetics of intellectual disability in consanguineous families. *Mol Psychiatry*. 2018. https://doi.org/10.1038/s41380-017-0012-2
19. Najmabadi H, Hu H, Garshasbi M, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*. 2011;478:57-63.
20. Reuter MS, Tawamie H, Buchert R, et al. Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. *JAMA Psychiat*. 2017;74:293-299.
21. Harripaul R, Vasli N, Mikhailov A, et al. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. *Mol Psychiatry*. 2018;23:973-984.
22. Martin HC, Jones WD, Stephenson D, et al. Quantifying the contribution of recessive coding variation to developmental disorders. *bioRxiv*. 2017. https://doi.org/10.1101/201533
23. Al-Mubarak B, Abouelhoda M, Omar A, et al. Whole exome sequencing reveals inherited and de novo variants in autism spectrum disorder: a trio study from Saudi families. *Sci Rep*. 2017;7:5679.
24. Bittles AH. Birth defect risk for children of first cousins is overstated. *The Conversation*. 2013;8. http://theconversation.com/birth-defect-risk-for-children-of-first-cousins-is-overstated-15809, accessed: August 04, 2018.
25. Donner S. *Mythos Verwandtenehe. Der Tagesspiegel*. Vol 23506. Berlin: Giovanni di Lorenzo, Sebastian Turner; 2018.
26. Hu H, Wienker TF, Musante L, et al. Integrated sequence analysis pipeline provides one-stop solution for identifying disease-causing mutations. *Hum Mutat*. 2014;35:1427-1435.
27. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet*. 2017;100:267-280.
28. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164.

29. McQuillan R, Leutenegger AL, Abdel-Rahman R, et al. Runs of homozygosity in European populations. *Am J Hum Genet*. 2008;83:359-372.
30. Kekatpure MV, Gaur S, Dash GK, Kannan S. TRPM6 mutation: a novel cause of "reversible" infantile epileptic encephalopathy. *Neurol India*. 2016;64:1037-1038.
31. Lesage S, Drouet V, Majounie E, et al. Loss of VPS13C function in autosomal-recessive parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent Mitophagy. *Am J Hum Genet*. 2016;98:500-513.
32. Esmaeeli-Nieh S, Fenckova M, Porter IM, et al. BOD1 is required for cognitive function in humans and drosophila. *PLoS Genet*. 2016;12:e1006022.
33. Karaca E, Harel T, Pehlivan D, et al. Genes that affect brain structure and function identified by rare variant analyses of Mendelian neurologic disease. *Neuron*. 2015;88:499-513.
34. Anazi S, Maddirevula S, Salpietro V, et al. Expanding the genetic heterogeneity of intellectual disability. *Hum Genet*. 2017;136:1419-1429.
35. Bras J, Verloes A, Schneider SA, Mole SE, Guerreiro RJ. Mutation of the parkinsonism gene ATP13A2 causes neuronal ceroid-lipofuscinosis. *Hum Mol Genet*. 2012;21:2646-2650.
36. Snijders Blok L, Hiatt SM, Bowling KM, et al. De novo mutations in MED13, a component of the mediator complex, are associated with a novel neurodevelopmental disorder. *Hum Genet*. 2018;137:375-388.
37. Ganna A, Genovese G, Howrigan DP, et al. Ultra-rare disruptive and damaging mutations influence educational attainment in the general population. *Nat Neurosci*. 2016;19:1563-1565.
38. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36:928-930.
39. Bittles AH, Neel JV. The costs of human inbreeding and their implications for variations at the DNA level. *Nat Genet*. 1994;8:117-121.
40. Hamamy H. Consanguineous marriages : preconception consultation in primary health care settings. *J Community Genet*. 2012;3:185-192.
41. Fareed M, Afzal M. Estimating the inbreeding depression on cognitive behavior: a population based study of child cohort. *PLoS One*. 2014;9:e109585.
42. Scott EM, Halees A, Itan Y, et al. Characterization of greater middle eastern genetic variation for enhanced disease gene discovery. *Nat Genet*. 2016;48:1071-1076.
43. Wahlund S. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*. 1928;11:65-106.
44. Overall ADJ. The influence of the wahlund effect on the consanguinity hypothesis: consequences for recessive disease incidence in a socially structured pakistani population. *Hum Hered*. 2009;67:140-144.
45. Bittles AH, Black ML. The impact of consanguinity on neonatal and infant health. *Early Hum Dev*. 2010;86:737-741.
46. Erkelenz S, Theiss S, Otte M, Widera M, Peter JO, Schaal H. Genomic HEXploring allows landscaping of novel potential splicing regulatory elements. *Nucleic Acids Res*. 2014;42:10681-10697.
47. Soukarieh O, Gaildrat P, Hamieh M, et al. Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. *PLoS Genet*. 2016;12:e1005756.
48. Soukarieh O, Gaildrat P, Hamieh M, et al. Correction: Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. *PLoS Genet*. 2016;12:e1005971.
49. Girirajan S, Rosenfeld JA, Coe BP, et al. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med*. 2012;367:1321-1331.
50. Girirajan S, Rosenfeld JA, Cooper GM, et al. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet*. 2010;42:203-209.
51. Heyne HO, Singh T, Stamberger H, et al. De novo variants in neurodevelopmental disorders with epilepsy. *Nat Genet*. 2018;50:1048-1053.
52. Heger M. UK's NHS Prepares Transition to Diagnostic Genome Sequencing for Rare Diseases, Some Cancers. http:/www.genomeweb.com/molecular-diagnostics/uks-nhs-prepares-transition-diagnostic-genome-sequencing-rare-diseases-some#.W6ytsX0YMhc: Genomeweb, 2018. accessed October 03, 2018.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Kahrizi K, Hu H, Hosseini M, et al. Effect of inbreeding on intellectual disability revisited by trio sequencing. *Clin Genet*. 2019;95:151–159. https://doi.org/10.1111/cge.13463