European Journal of Human Genetics (2011) 19, 1161–1166

www.nature.com/ejhg

ARTICLE

Homozygosity mapping in 64 Syrian consanguineous families with non-specific intellectual disability reveals 11 novel loci and high heterogeneity

Rami Abou Jamra*[,1,2], Sigrun Wohlfart[1], Markus Zweier[1], Steffen Uebe[1], Lutz Priebe[2,3], Arif Ekici[1], Susanne Giesebrecht[2], Ahmad Abboud[4], Mohammed Ayman Al Khateeb[5], Mahmoud Fakher[6], Saber Hamdan[7], Amina Ismael[8], Safia Muhammad[9], Markus M Nöthen[2], Johannes Schumacher[2] and André Reis[1]

Non-specific intellectual disability of autosomal recessive inheritance (NS-ARID) represents an important fraction of severe cognitive dysfunction disorders. To date, only 10 genes have been identified, and further 24 linked-ARID loci have been reported, as well as others with suggestive linkage. To discover novel genes causing NS-ARID, we undertook genome-wide homozygosity mapping in 64 consanguineous multiplex families of Syrian descent. A total of 11 families revealed unique, significantly linked loci at 4q26-4q28 (MRT17), 6q12-q15 (MRT18), 18p11 (MRT19), 16p12-q12 (MRT20), 11p15 (MRT21), 11p13-q14 (MRT23), 6p12 (MRT24), 12q13-q15 (MRT25), 14q11-q12 (MRT26), 15q23-q26 (MRT27), and 6q26-q27 (MRT28), respectively. Loci ranged between 1.2 and 45.6 Mb in length. One family showed linkage to chromosome 8q24.3, and we identified a mutation in *TRAPPC9*. Our study further highlights the extreme heterogeneity of NS-ARID, and suggests that no major disease gene is to be expected, at least in this study group. Systematic analysis of large numbers of affected families, as presented here, will help discovering the genetic causes of ID.
European Journal of Human Genetics (2011) **19,** 1161–1166; doi:10.1038/ejhg.2011.98; published online 1 June 2011

Keywords: non-specific; mental retardation; microcephaly; epilepsy; MRT

INTRODUCTION

Intellectual disability (ID) is the most common reason for referral to clinical genetic centers, and is an unsolved problem in health care with important impact on the individuals, their families, and the health system. Although mild forms of ID (intelligence quotient (IQ) 70–50) may, in some cases, be of genetic complex nature, severe forms (IQ < 50, incidence of 0.4%) are thought to be caused by striking environmental or genetic factors.[1] In addition to the classification based on severity, patients may be classified into syndromic and non-syndromic forms. In syndromic forms, dysmorphisms, biochemical parameters, imaging findings, and organ anomalies exist, and may help identifying a syndrome of known etiology. In the non-syndromic form, further non-specific symptoms such as epilepsy, ataxia, mild microcephaly, and behavioral disorders may exist. This means that a characteristic pattern of subtle symptoms may be overlooked in the first patient presenting, and will be recognized only when examining several patients with mutations in the same gene. Non-specific ID may thus be considered as an exclusion diagnosis of distinguishable, syndromic ID, with a considerable gray zone between both forms.

In the last few years, it became clear that genetic factors, that is, chromosomal anomalies and point mutations, have a central role in the etiology of ID. In a representative study, Rauch *et al*[2] showed that 15% of the cases of ID are due to cytogenetic anomalies, and that further 15% are due to submicroscopic aberrations. X-chromosomal mutations are possibly the cause in about 10% of the cases.[3] This means that in more than half of cases, the etiology of ID remains unsolved. The majority of these have probably an autosomal form of ID. In industrial countries, both dominant and recessive forms usually appear as sporadic cases due to small family size. This hampers genetic approaches for identifying the underlying genetic cause. Genetically distinct subtypes are often clinically indistinguishable, precluding pooling of data from unrelated families, and rendering their elucidation even more difficult. It is therefore not surprising, that to date, only few genes have been identified (recently reviewed by Ropers).[1]

The frequency of autosomal recessive, non-specific ID is unknown. Although functional considerations and empirical data from mouse models suggest that most gene defects are inherited as recessive disorders,[1,4,5] a recent study suggests that autosomal dominant *de novo* mutations are most prevalent.[6] Until now, 26 significantly linked non-specific intellectual disability of autosomal recessive inheritance (NS-ARID) loci were described,[7–10] and only 10 genes (six of them located in described regions) were identified: *PRSS12* (OMIM#606709, on 4q26, former MRT1), *CRBN* (OMIM#609262, on 3p26, former MRT2), *CC2D1A* (OMIM#610055, on 19p13.12, former MRT3), *ST3GAL3* (OMIM#606494, on 1p34.3), *GRIK2* (OMIM#138277, on 6q16, former MRT6), *TUSC3* (OMIM#601385,

[1]Institute of Human Genetics, University of Erlangen, Erlangen, Germany; [2]Institute of Human Genetics, University of Bonn, Bonn, Germany; [3]Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany; [4]Practising Paediatrician in Jdaidet Artouz, Syria; [5]Practising Paediatrician in Sanameen, Syria; [6]Practising Paediatrician in Shahbaa, Syria; [7]Practising Paediatrician in Kraya, Syria; [8]Practising Paediatrician in Jesser El Sheghour, Syria; [9]Practising Paediatrician in Lattakia, Syria
*Correspondence: Dr R Abou Jamra, Institute of Human Genetics, University of Erlangen-Nuremberg, Schwabachanlage 10, 91054 Erlangen, Germany. Tel: +49 9131 85 26477; Fax: +49 9131 85 23232; E-mail: rami.aboujamra@uk-erlangen.de
Received 15 November 2010; revised 1 March 2011; accepted 19 April 2011; published online 1 June 2011

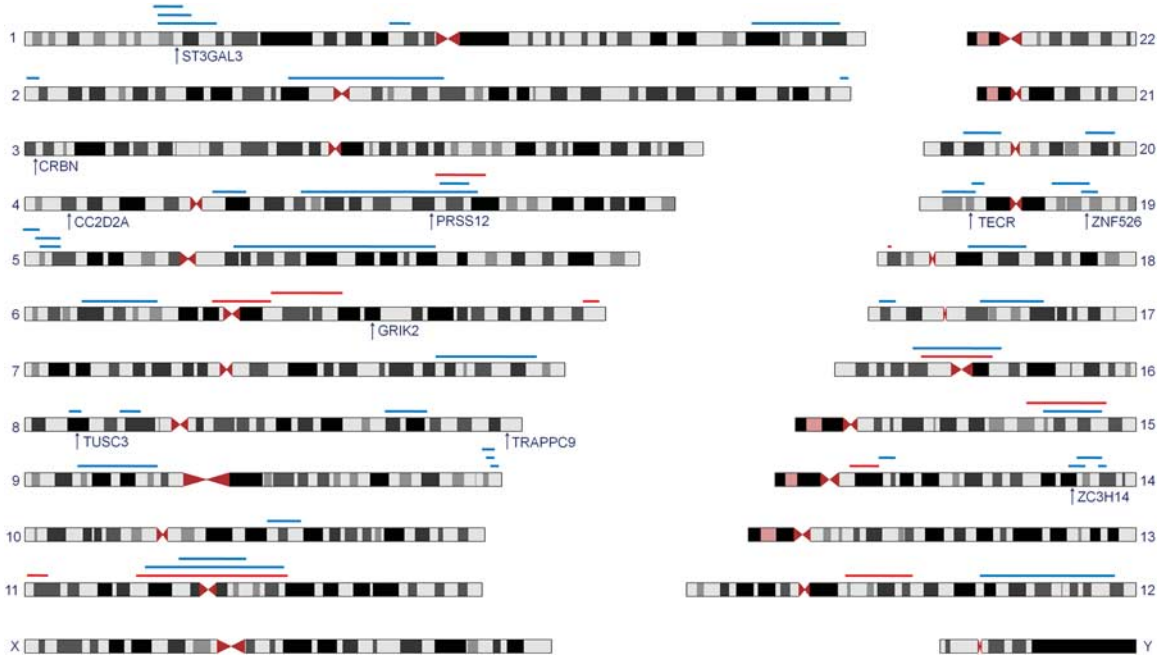

Figure 1 Overview of NS-ARID loci reported to date[7–10] (blue bars), and novel loci reported here (red bars). Positions of known NS-ARID genes are marked with an arrow next to the gene symbol. Diagram was drawn with help of the GenomeGraph tool of the UCSC genome browser.

on 8p22, former MRT7), *ZNF26* (no OMIM#, on 19q13.2), *TRAPPC9* (OMIM#613192, on 8q24, former MRT13), *ZCH14* (OMIM# 613279, on 14q31.3), and *TECR* (MIM*610057 on 19p13.12) (Figure 1).[1,11–19]

Taking into account that over 90 genes are responsible for only about 40% of X-chromosomal recessive ID cases, and that about half of the estimated 22 000 human genes are expressed in the brain, the total number of ARID genes may run into the hundreds.[1] Considering the heterogeneity observed in previous linkage studies, and the fact that the so far identified genes do not account for a significant fraction of NS-ARID cases, systematic approaches are the most promising strategy to identify further genes.[7,8,10] Large consanguineous families provide enough information to map loci, based on analysis in a single family and thus represent the best starting point. We present here the results of the homozygosity mapping in a series of 64 Syrian multiplex consanguineous families with NS-ARID.

PATIENTS AND METHODS

Affected families were identified with the help of local pediatricians. After obtaining written informed consent, all affected individuals were clinically evaluated by a pediatrician (AAb, AAl, MF, SH, AI, or SM) and by a geneticist (RAJ), with special attention to neurological, morphological, ophthalmological, dermatological, and skeletal symptoms. The mental status of affected and unaffected individuals was estimated by investigation of verbal and motor aptitudes at the time of visit. Parents were interviewed to discriminate facts about prenatal, perinatal, and neonatal medical history of all children, either affected or unaffected. As formal assessment of ID was not possible, we used a description of developmental milestones and abilities to evaluate severity of ID, as previously described.[20] A detailed pedigree was constructed, and the genealogical relationship within families was crosschecked by interviewing different family members. Blood probes were drawn, and DNA was isolated from all available individuals. For one patient of each nuclear family, testing for Fragile X syndrome was carried out by PCR. In addition, all mothers were checked for X-inactivation status.

From the 85 families initially ascertained, 11 were excluded either for not providing sufficient blood samples, non-random X-inactivation in the mother of the index case suggesting X-linked inheritance, or non-informativeness of clinical and familial history. Four further families were excluded, because a diagnosis could be made (phenylketonuria, Miller-Dieker-Syndrome, van der Knaap Leukoencephalopathy, Cohen syndrome). Six further families with only one affected child each were also not analyzed further, as an autosomal dominant cause could not be excluded. Of the remaining 64 families, six nuclear families could be aggregated into two complex families, with three branches of affected cousins each, and further 16 families into eight complex families, with two branches each. In total, 41 nuclear and 10 complex families consisted of 305 children (140 affected), and 128 parents were subject to molecular genetics analysis. Of those, 31 families had two affected children, and 20 families had three or more affected children. Kinship co-efficiencies varied between 0.0039 and 0.1563. Of these, 12 families had only affected males, and 11 only affected females. Though X-chromosomal recessive inheritance cannot be definitely excluded, this is unlikely, as mothers of affected males were healthy, had several healthy brothers, nephews, and uncles on the maternal side, Fragile X syndrome was excluded, and X-inactivation was unremarkable.

Affected children presented either with moderate ID in 45, or severe ID in 19 nuclear families. All cases were non-specific, that is, no peculiar or distinguishable combination of symptoms was found. Further, unspecific symptoms we documented were epilepsy in 17 families, growth retardation in 11 families, muscular hypertonia in five families and hypotonia in 17 families, microcephaly in 14 families, ataxia in four families, and developmental regression in three families. The symptoms and severity were comparable in sibs. As many syndromes have variable phenotypes, we cannot exclude syndromic forms with certainty (Tables 1).

All available affected individuals and their healthy siblings were genotyped, healthy parents only when grandparents were also consanguineous. In total, 154 healthy and 128 affected persons were genotyped, 11 using Human610-Quad DNA Analysis BeadChips (Illumina, San Diego, CA, USA), 72 with HumanCytoSNP-12 DNA Analysis BeadChips (Illumina), and 201 with Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) microarrays. The platform used was consistent within each family, with the exception of three individuals recruited at a later stage. Marker density was approximately 600 000, 300 000, and 900 000 SNPs, respectively. The minimal coverage of approximately 300 000 SNPs is more than sufficient to identify homozygosities of at least 500 kb (average of 50 markers).

Table 1 Structure and clinical description of the 12 families, which showed one linkage locus

Family	*# Aff*[a]	*Kinship co-efficient*[b]	*Age at presentation of affected individuals (years)*	*Symptoms*[c]
MR003	5fm	0.097	18, 14, 9, 6, 2	Neonatal muscular hypotonia, mild motor delay, severe ID, no speech, understanding better, no autistic symptoms, MRI and further detailed examinations unremarkable
MR006	4f	0.023, 0.094	15, 10, 7, 2	Neonatal muscular hypotonia, severe motor delay, no walking, crawling, muscular hypotonia, severe ID, no speech, weak interaction with surrounding, macroglossia, growth retardation and mild microcephaly, scream loud and often
MR013	3fm	0.133, 0.016	43, 23, 19	Mild motor delay, moderate ID, can speak and interact with surrounding properly, arthrogryposis of several small joints, deformity of feet in two affected, no or discrete facial gestalt, amicable, cheerful
MR019	4fm	0.004, 0.005	6, 2, 2, 1	Mild motor delay, moderate ID, can speak single words, understanding better, no or discrete facial gestalt, early onset of epilepsy by three patients
MR043	3f	0.063, 0.078	8, 6, 4	Severe motor delay, no walking, no sitting or sitting at age of 5 yrs, severe ID, no or single words, absence epilepsy by one patient, mild microcephaly
MR050	2m	0.063	9, 4	Neonatal muscular hypotonia, later hypertonia of the upper extremities by one patient, moderate motor delay, severe ID, tremor, no speech, understanding better, CT unremarkable
MR055ab	6fm	0.078, 0.008	33, 23, 23, 13, 7, 7	Neonatal muscular hypotonia, severe motor delay, single words or no speech, simple understanding or no understanding, severe or mild growth retardation, severe or moderate microcephaly, stereotypic movements, hand-flapping, facial gestalt
MR056	3fm	0.004	17, 12, 11	Microcephaly and small size at birth, motor delay, delayed and simple speech, good understanding, moderate ID, growth retardation, mild microcephaly, MRI unremarkable
MR059	3fm	0.078	11, 10, 8	Mild motor delay, delay of speech, moderate ID, vision impairment (not defined), autistic symptoms, MRI unremarkable, anxiety
MR061bc	5fm	0.016, 0.063	30*, 22, 20, 18*, 15*	Severe motor delay, walked for few months and then lost ability to walk, atrophy and spastic of lower extremeties, severe ID, no speech, weak interaction with surrounding, normal or mild microcephaly, growth retardation
MR068	5fm	0.063	19, 18, 15, 7, 6	Normal motor development, delayed speech, moderate ID, epilepsy
MR073	3m	0.094	10, 7, 1.5	Neonatal normal development, moderate ID, mild microcephaly, moderate motor delay, severe lingual delay, CT shows deficiencies of the white matter

[a]Number of affected individuals and their gender, for example, 5fm means five affected individuals of both sexes.
[b]Degree of relatedness between parents, for example, 0.063 for I° cousins. In complex families, values for each nuclear family are given. These values are lower estimates, as often relationships are more complex than could be reliably documented.
[c]If not otherwise indicated, pregnancy and birth were unremarkable, no facial gestalt, no epilepsy, no growth retardation, and no microcephaly were observed.
*These patients were dead at the time of examination.

To detect the presence of copy number variants, molecular karyotyping was performed using either the Affymetrix Genotyping Console Software (Version 3.0.2), or the PennCNV software[21] for the Illumina arrays. In identical by descent (IBD) regions, we set filter criteria to $\geq$10 kb and $\geq$5 markers (Affymetrix), and $\geq$20 kb and a confidence value $\geq$10 (Illumina). To discriminate between genomic polymorphisms and aberrations containing MR-associated genes or regions, we compared findings within each family, with the database of Genomic Variants (http://projects.tcag.ca/variation/) and with 820 (Affymetrix) and 750 (Illumina) internal controls.

Mendelian segregation was calculated using PedCheck and the EasyLinkage interface software, and was confirmed in all instances.[22–24] Genome-wide analysis was performed using Homozygosity Mapper and re-checked manually.[25] In addition, linkage analysis was performed using ALLEGRO or GeneHunter under an autosomal recessive mode of inheritance, with 99% penetrance and a disease allele frequency of 0.001, using the EasyLinkage interface software.[23,26,27]

RESULTS AND DISCUSSION

Novel ARID gene loci

Twelve families showed only one IBD region each, with significant LOD scores above 3. Eleven of these were not included in published regions and were thus designated MRT numbers (HUGO Gene Nomenclature Committee, http://www.genenames.org/index.html, see Tables 1 and 2). In the 12th family, MR055, we identified a mutation in *TRAPPC9*. The linked regions vary in length between 1.2 and 45.6 Mb, and include between 9 and 625 RefSeq genes (Table 2).

In the era of Next-Generation Sequencing, families with chromosomal regions IBD are of great value, even if they have more than one IBD region. In addition to the 11 families with a single locus, further 17 families showed IBD regions with a total length of less than 1% of the genome ($<$31 Mb), distributed over two (six families), three (eight families), four (one family), or five (two families) IBD regions. These IBD segments will be highly valuable for identifying further mutations in candidate genes. CNV analysis in IBD regions identified no deletion or duplication in candidate genes. In one family with seven IBD regions, we found a relatively common, recently described deletion over the gene *OTOA* at 16p12.2, explaining their hearing impairment, but not the ID.[28]

Identifying the causing mutation in *TRAPPC9* in family MR055

Family MR055 consists of two branches (a and b). In the first branch, parents were cousins I°, and there were three affected females (two are true twins) and one healthy male. In the second branch, parents were cousins III° once removed, and there were three affected males (two are true twins) and one healthy female. Both families are distantly related, as great-grand parents were cousins.

At the time of examination, the three affected boys of family MR055a were 13 and 7 (twins) years old, 141, 116, and 116 cm tall, respectively, and had head circumferences of 47, 45.5, and 46 cm, respectively. Pregnancies and births were unremarkable; in the neonatal period, parents reported muscular hypotonia. The affected boys

Table 2 Genetic findings in 12 families showing a unique linkage locus

		Locus (in Mb)									
Family		*Start SNP*	*End SNP*	*Chr.*	*Start*	*End*	*Length*	*Cyto band*	*Peak LOD-score*	*No. RefSeq genes*	*Array used*
MR003	MRT17	rs10518325	rs10518608	4	119.9	133.3	13.3	4q26–q28	3.33	43	610K
MR006	MRT18	rs4612125	rs285651	6	69.1	91.1	22.0	6q12–q15	4.55	93	CytoSNP
MR013	MRT19	rs4606805	rs1787846	18	2.4	3.6	1.2	18p11	3.03	13	CytoSNP
MR019	MRT20	rs7197568	rs7197227	16	24.0	46.7	21.7	16p12–q12	3.91	156	CytoSNP
MR043	MRT21	rs3802985	rs7126612	11	0.2	6.6	6.4	11p15	3.55	193	Affy6.0
MR050	MRT23	rs604518	rs10899421	11	31.8	77.4	45.6	11p13–q14	3.03	625	Affy6.0
MR055	*TRAPPC9*	rs6577946	rs3942977	8	139.8	142.6	2.8	8q24	4.67	12	Affy6.0
MR056	MRT24	rs651733	rs1508668	6	52.9	69.3	16.4	6p12	3.39	31	Affy6.0
MR059	MRT25	rs4760658	rs1882033	12	46.6	66.1	19.5	12q13–q15	3.36	323	Affy6.0
MR061	MRT26	rs10132585	rs1278951	14	22.4	31.5	9.1	14q11–q12	3.85	90	Affy6.0
MR068	MRT27	rs868127	rs2388310	15	68.2	91.8	23.6	15q23–q26	3.38	236	Affy6.0
MR073	MRT28	rs6935718	rs3886091	6	162.2	166.6	4.4	6q26–q27	3.36	9	Affy6.0

In family MR055, a nonsense mutation in *TRAPPC9* was identified (for details see text).

learned to walk at age of 5 years; they can understand simple phrases, but cannot speak. All have stereotypic movements and hand-flapping behavior. The elder brother had few epileptic seizures. The affected brothers resemble each other with low frontal hairline, synophrys, and microcephaly. At the time of examination, the three affected women of family MR055b were 33 and 23 (twins) years old, 132, 133, and 139 cm tall, respectively, and had head circumferences of 43.5, 42, and 42 cm, respectively. Pregnancies and births were unremarkable. They learned to walk at age of 7 years, never learned speaking, and they do not understand simple phrases. All have stereotypic movements and hand-flapping behavior. None of the affected had epileptic seizures. The elder affected woman is spontaneously losing her teeth and losing weight. The parents reported that this started at age of 20, which can be observed by the younger two affected sisters (23 years at the time of examination). As in the first branch, the affected sisters resemble each other with low frontal hairline, synophrys, and microcephaly.

Family MR055 showed significant linkage on chromosome 8 (Table 2), including *TRAPPC9*, which is a perfect candidate gene.[16–18] We sequenced the gene and identified the mutation c.1423C>T;p.R475X recently reported in a Palestinian family.[17] Because of the geographical and ethnical proximity (our family MR055 is from the south of Syria), and assuming an extremely low frequency of this mutation, we postulate that this is a founder mutation of this specific population and/or geographical area. Symptoms of the Palestinian and Syrian patients were comparable, moderate to severe ID, microcephaly, hand-flapping movements, no regression, and no epilepsy. In contrast to our patients, Mir *et al*[16] reported normal height and less severe ID, and Philippe *et al*[18] reported normal height and obesity. Thus, we conclude that mutations in *TRAPPC9* cause IDs, with tendency to severe forms and microcephaly. Further symptoms such as short stature, stereotypic movements, epilepsy, and remarkable MRI findings seem to be not specific. Because of the relative similarity of our and the Palestinian patients, we do not exclude a genotype–phenotype correlation.

Loci overlapping known ARID loci

Apart from family MR055 with the *TRAPPC9* mutation, none of the families with one linkage locus overlapped any of the 10 to-date-described ARID genes. Nonetheless, five families overlap seven already known loci, but with yet undetected causative gene (Figure 1). Overlapping findings are of importance, as these might facilitate the finding of a second allele. Further, we checked for overlapping IBD regions across all families studied, that is, also families with several IBDs. We identified five regions, each with overlapping IBDs in five families on chromosomes 3 (175.5–177.5 and 178.5–180.6 Mb), 19 (7.6–7.9 and 8.3–9.2 Mb), and 22 (20.2–21.1 Mb, all hg19) and 13 regions with four overlapping IBDs, each. These regions harbor interesting candidates for further analysis. Nevertheless, in the 51 families, we identified a total of 275 IBD segments encompassing 3196 Mb; thus any overlap may also be coincidental.

Intrafamilial heterogeneity in complex families

The two complex families MR061 and MR071, each consisting of three branches, a, b, and c, are an example of the impressive heterogeneity of the NS-ARID phenotype (Figure 2). Both families did not show one common linkage region for all three branches. When analyzed together, families MR061a and MR061b showed one homozygous locus on chromosome 14q11-q12 (MRT26, 9.1 Mb long, LOD-score of 3.85), whereas branch c does not map to this locus (Figure 2). The nuclear families MR071a, MR071b, and MR071c also did not show a common homozygote region. Depending on the analyzed combination of nuclear families, different linked loci are possible. Although the combination of either MR071a and MR071b, or MR071a and MR071c showed four linkage regions of up to 30 Mb with LOD scores of ≥ 2.4 each, the combination MR071b and MR071c showed one homozygote region of about 33 Mb on chromosome 15q21-q25 (LOD-score of 2.7). Though we would expect that genotyping further members of this family of extensive consanguinity (kinship co-efficient >0.15) might define the true linkage region, this family shows that even high LOD-scores must be considered cautiously. On the other hand, suggestive LOD scores may turn out to be true findings. An example has been shown by Najmabadi *et al*,[7] who identified a LOD score of only two in family M001, and later identified a mutation in *TRAPPC9*, which is located in this region and proved that this locus is true, though the linkage analysis was not significant.[7,16]

Eight complex families MR006, MR013, MR018, MR019, MR022, MR043, MR049, and MR055, have two branches each, a, and b. Families MR022a and MR022b, and MR049a and MR049b showed no common homozygous region (Supplementary Figure 1). Five of the other six families, MR006, MR013, MR019, MR043, and MR055

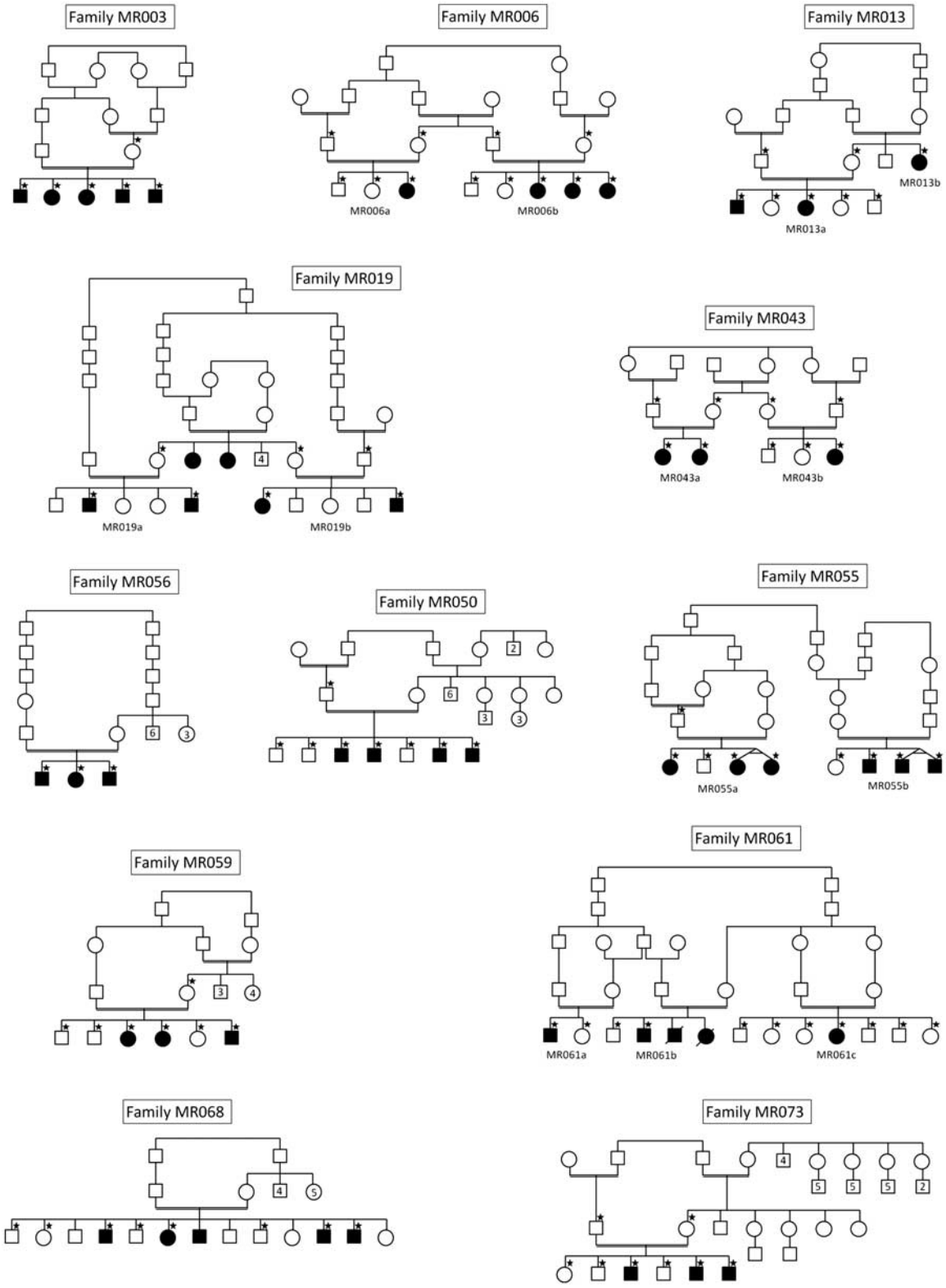

Figure 2 Overview of the 12 families, which showed one linkage locus. The families were simplified as possible, and important information such as healthy brothers of mothers or grandmothers are shown. Asterisks (*) indicate genotyped DNAs.

showed one common homozygous region with LOD scores >3, each, on 6q12-q15 (MRT18), 18p11.32-p11.31 (MRT19), 16p12.1-q12.1 (MRT20), 11p15.5-15-4 (MRT21), and over *TRAPPC9*, respectively (Table 2).

We conclude that large and complex families represent a promising template to identify candidate regions, but bear the risk of false positive results because of heterogeneity. At the extreme end, such heterogeneity may exist within the same nuclear family.

CONCLUSIONS

Our findings suggest that NS-ARID is extremely heterogeneous in the Syrian population, which is in line with earlier findings in other populations.[7,8] Autosomal recessive disorders are even a greater challenge in populations with low consanguineous marriages, as those will often be compound heterozygote, and thus not due to IBD mutations. In addition, it seems certain by now that mutations in the currently known NS-ARID genes (*PRSS12*, *CRBN*, *CC2D1A*, *GRIK2*, *TUSC3*, *TRAPPC9*, *ZNF526*, *ZC3H14*, *ST3GAL3*, and *TECR*) are responsible for only a small fraction of NS-ARID cases, whereas hundreds of additional genes are likely to be identified in the near future. The most promising concept for this is the systematic analysis of a large number of consanguineous families, followed by ultra deep sequencing strategies. Identification of these genes will improve genetic counseling of affected families, and advance our understanding of molecular networks involved in cognitive processes. Understanding these processes should allow the replacement of symptomatic and supportive therapies, with pharmacotherapies based on a principled understanding of the causes of cognitive impairment as is already becoming evident for several neurodevelopmental disorders.[29]

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the patients and their families for their cooperation. We also thank Angelika Diem, Nadine Kluck, and Petra Rothe for the excellent technical support. We thank Dr Cornelia Kraus for the diagnostics of Fra X syndrome and X-inactivation. R Abou Jamra was supported by the *Deutsche Forschungsgemeinschaft* (DFG) (grant AB393/1-1 and 1-2). This study was supported by the German Intellectual disability Network (MRNET) through a grant from the German Ministry of Research and Education to A Reis (01GS08160).

1 Ropers HH: Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 2010; **11**: 161–187.
2 Rauch A, Hoyer J, Guth S *et al*: Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* 2006; **140**: 2063–2074.
3 Ropers HH, Hamel BC: X-linked mental retardation. *Nat Rev Genet* 2005; **6**: 46–57.
4 Bartley JA, Hall BD: Mental retardation and multiple congenital anomalies of unknown etiology: frequency of occurrence in similarly affected sibs of the proband. *Birth Defects Orig Artic Ser* 1978; **14**: 127–137.
5 Priest JH, Thuline HC, Laveck GD, Jarvis DB: An approach to genetic factors in mental retardation. Studies of families containing at least two siblings admitted to a state institution for the retarded. *Am J Ment Defic* 1961; **66**: 42–50.
6 Vissers LE, de Ligt J, Gilissen C *et al*: A *de novo* paradigm for mental retardation. *Nat Genet* 2010; **42**: 1109–1112.
7 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.
8 Rafiq MA, Ansar M, Marshall CR *et al*: Mapping of three novel loci for non-syndromic autosomal recessive mental retardation (NS-ARMR) in consanguineous families from Pakistan. *Clin Genet* 2010; **78**: 478–483.
9 Uyguner O, Kayserili H, Li Y *et al*: A new locus for autosomal recessive non-syndromic mental retardation maps to 1p21.1-p13.3. *Clin Genet* 2007; **71**: 212–219.
10 Kuss AW, Garshasbi M, Kahrizi K *et al*: Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum Genet* 2011; **129**: 141–148.
11 Molinari F, Rio M, Meskenaite V *et al*: Truncating neurotrypsin mutation in autosomal recessive nonsyndromic mental retardation. *Science* 2002; **298**: 1779–1781.
12 Higgins JJ, Pucilowska J, Lombardi RQ, Rooney JP: A mutation in a novel ATP-dependent Lon protease gene in a kindred with mild mental retardation. *Neurology* 2004; **63**: 1927–1931.
13 Motazacker MM, Rost BR, Hucho T *et al*: A defect in the ionotropic glutamate receptor 6 gene (GRIK2) is associated with autosomal recessive mental retardation. *Am J Hum Genet* 2007; **81**: 792–798.
14 Garshasbi M, Hadavi V, Habibi H *et al*: A defect in the TUSC3 gene is associated with autosomal recessive mental retardation. *Am J Hum Genet* 2008; **82**: 1158–1164.
15 Molinari F, Foulquier F, Tarpey PS *et al*: Oligosaccharyltransferase-subunit mutations in nonsyndromic mental retardation. *Am J Hum Genet* 2008; **82**: 1150–1157.
16 Mir A, Kaufman L, Noor A *et al*: Identification of mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation. *Am J Hum Genet* 2009; **85**: 909–915.
17 Mochida GH, Mahajnah M, Hill AD *et al*: A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. *Am J Hum Genet* 2009; **85**: 897–902.
18 Philippe O, Rio M, Carioux A *et al*: Combination of linkage mapping and microarray-expression analysis identifies NF-kappaB signaling defect as a cause of autosomal-recessive mental retardation. *Am J Hum Genet* 2009; **85**: 903–908.
19 Caliskan M, Chong JX, Uricchio L *et al*: Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. *Hum Mol Genet* 2011; **20**: 1285–1289.
20 Zhang X, Snijders A, Segraves R *et al*: High-resolution mapping of genotype-phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. *Am J Hum Genet* 2005; **76**: 312–326.
21 Wang K, Li M, Hadley D *et al*: PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 2007; **17**: 1665–1674.
22 O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
23 Hoffmann K, Lindner TH: EasyLINKAGE-Plus–automated linkage analyses using large-scale SNP data. *Bioinformatics* 2005; **21**: 3565–3567.
24 Lindner TH, Hoffmann K: EasyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses. *Bioinformatics* 2005; **21**: 405–407.
25 Seelow D, Schuelke M, Hildebrandt F, Nurnberg P: HomozygosityMapper–an interactive approach to homozygosity mapping. *Nucleic Acids Res* 2009; **37**: W593–W599.
26 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
27 Gudbjartsson DF, Thorvaldsson T, Kong A, Gunnarsson G, Ingolfsdottir A: Allegro version 2. *Nat Genet* 2005; **37**: 1015–1016.
28 Shahin H, Walsh T, Rayyan AA *et al*: Five novel loci for inherited hearing loss mapped by SNP-based homozygosity profiles in Palestinian families. *Eur J Hum Genet* 2010; **18**: 407–413.
29 Wetmore DZ, Garner CC: Emerging pharmacotherapies for neurodevelopmental disorders. *J Dev Behav Pediatr* 2010; **31**: 564–581.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)