

Clin Genet 2012: 82: 248–255 Printed in Singapore. All rights reserved

Short Report



© 2011 John Wiley & Sons A/S CLINICAL GENETICS doi: 10.1111/j.1399-0004.2011.01755.x

Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of *ARID1B*

Halgren C, Kjaergaard S, Bak M, Hansen C, El-Schich Z, Anderson CM, Henriksen KF, Hjalgrim H, Kirchhoff M, Bijlsma EK, Nielsen M, den Hollander NS, Ruivenkamp CAL, Isidor B, Le Caignec C, Zannolli R, Mucciolo M, Renieri A, Mari F, Anderlid B-M, Andrieux J, Dieux A, Tommerup N, Bache I. Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of *ARID1B*.

Clin Genet 2012: 82: 248-255. © John Wiley & Sons A/S, 2011

Corpus callosum abnormalities are common brain malformations with a wide clinical spectrum ranging from severe intellectual disability to normal cognitive function. The etiology is expected to be genetic in as much as 30-50% of the cases, but the underlying genetic cause remains unknown in the majority of cases. By next-generation mate-pair sequencing we mapped the chromosomal breakpoints of a patient with a de novo balanced translocation, t(1;6)(p31;q25), agenesis of corpus callosum (CC), intellectual disability, severe speech impairment, and autism. The chromosome 6 breakpoint truncated ARID1B which was also truncated in a recently published translocation patient with a similar phenotype. Quantitative polymerase chain reaction (Q-PCR) data showed that a primer set proximal to the translocation showed increased expression of ARID1B, whereas primer sets spanning or distal to the translocation showed decreased expression in the patient relative to a non-related control set. Phenotype-genotype comparison of the translocation patient to seven unpublished patients with various sized deletions encompassing ARID1B confirms that haploinsufficiency of ARID1B is associated with CC abnormalities, intellectual disability, severe speech impairment, and autism. Our findings emphasize that ARID1B is important in human brain development and function in general, and in the development of CC and in speech development in particular.

Conflict of interest

The authors declare no conflict of interests.

Re-use of this article is permitted in accordance with the Terms and Conditions set out at http://wileyonlinelibrary.com/onlineopen#Online Open_Terms

C Halgren^a, S Kjaergaard^b, M Bak^a, C Hansen^a, Z El-Schich^a, CM Anderson^a, KF Henriksen^a, H Hjalgrim^c, M Kirchhoff^b, EK Bijlsma^d, M Nielsen^d, NS den Hollander^d, CAL Ruivenkamp^d, B Isidor^e, C Le Caignec^e, R Zannolli^f, M Mucciolo^g, A Renieri^g, F Mari^g, B-M Anderlid^h, J Andrieuxⁱ, A Dieux^j, N Tommerup^a and I Bache^a

^aWilhelm Johannsen Centre for Functional Genome Research. Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark, ^bDepartment of Clinical Genetics, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark, ^cKlinik for Børn, Copenhagen, Denmark, ^dDepartment of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands, ^eCHU Nantes, Service de Génétique Médicale, Nantes, France, ^fDepartment of Pediatrics, and ^gMedical Genetics, Department of Biotechnology, University of Siena, Siena, Italy, hDepartment of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden, ⁱInstitut de Génétique Médicale, Hopital Jeanne de Flandre, CHRU de Lille, Lille, France, and ^jClinique de Génétique Médicale, CHRU de Lille, Lille, France

Key words: *ARID1B* – autism spectrum disorder – chromosome 6q25 – corpus callosum – intellectual disability – next-generation mate-pair sequencing – speech impairment – translocation

Corresponding author: Christina Halgren, Wilhelm Johannsen Centre for Functional Genome Reseach, Department of Cellular and Molecular Medicine, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej

3, 24.4.14, 2200 Copenhagen, Denmark. Tel.: +45 3532 7818; fax: +45 3532 7845; e-mail: halgren@sund.ku.dk

Received 19 May 2011, revised and accepted for publication 25 July 2011

The corpus callosum (CC) is the main interhemispheric commissure transferring cognitive, sensory, and motor information between the two brain hemispheres. CC abnormalities include complete agenesis, hypoplasia, and varied degrees of partial agenesis (1). Agenesis of CC (ACC) occurred in 1 in 1000 in a series of unselected neonates (2) and is thus one of the most common brain malformations. It is a heterogeneous condition with a wide clinical spectrum ranging from severe intellectual disability to normal cognitive function (3, 4). The etiology is believed to be genetic in 30-50% of the cases (5, 6) whereas fetal infections and exposure to teratogenes, e.g. alcohol, are suspected causes in the remaining cases. Numerous chromosomal loci have been associated with ACC (7, 8) including loci at 6q25q27 (8–11), but the underlying genetic cause remains unknown in the majority of cases.

Here we report eight previously unpublished patients with haploinsufficiency of *ARID1B*: one patient with a *de novo* translocation t(1;6)(p31;q25) mapped by next-generation sequencing (NGS) and seven patients with various sized *de novo* deletions.

Materials and methods

Patients

Each patient was clinically and molecularly evaluated by at least one of the authors. Patient 1 was identified through a national study of carriers of structural rearrangements; the study was approved by the Danish Scientific Ethics Committee and the Danish Data Protection Agency and written informed consent was obtained. Patients 2–8 were referred to genetic evaluation due to developmental delay; informed consent was obtained at the local clinical genetics departments. Patients 3–8 were identified in DECIPHER (12).

Chromosome analysis

Standard G-banding chromosome analysis was performed on cultured peripheral lymphocytes.

Next-generation paired-end sequencing

Mate-pair libraries were prepared using the Mate Pair Library v2 kit (Illumina, San Diego, CA). Briefly, 10 μ g genomic DNA was sheared using a Nebulizer. Fragments of 2–3 kb were isolated, end-repaired using a mix of natural and biotinylated dNTPs, blunt-end ligated using circularization ligase, and fragmented to 200-400 bp. Biotinylated fragments were isolated and end-repaired and A-overhangs were added to the 3' ends. Paired-end adapters were ligated to the fragments and the library was amplified by 18 cycles of PCR. Mate-pair libraries were subjected to $2 \times$ 36 bases paired-end sequencing on a Genome Analyzer IIx (Illumina), following the manufacturers protocol. Reads were aligned to a reference genome using Bowtie (13) allowing up to two mismatches in the seed region. Reads not aligning uniquely were discarded from further analysis. Paired reads aligning to different chromosomes or with unexpected strand orientation were extracted to identify potential translocation and inversion breakpoints, respectively. Breakpoints were only considered as candidates if they were confirmed by at least three independent paired reads with endreads mapping within a 6 kb region. Predicted breakpoints were filtered against known in-house variants based on data from 30 individuals with known breakpoints. Breakpoints were confirmed by PCR amplification and Sanger sequencing of the breakpoint-spanning fragments.

Quantitative polymerase chain reaction

RNA from patient 1 and five controls was extracted from peripheral blood using standard procedures. Following extraction, RNA was DNAse I (Invitrogen, San Diego, CA) treated and reverse transcribed with a HT11V primer using SuperscriptII (Invitrogen). Primers for *ARID1B* were designed using OLIGO software (Molecular Biology Insights Inc., W. Cascade, CO) (Table S1, Supporting information). All primer sets were designed to span at least one intron. Q-PCR was performed on an Opticon3 thermocycler (Bio-Rad Laboratories, Hercules, CA). All samples were run in triplicates. Normalization of expression was done using two stable housekeeping genes (*EIF6* and *G6PD*). Assessment of stable housekeeping genes was done using GENORM software (14).

Microarray analysis

Patient 1 was examined with Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA). copy number variations (CNVs) >1 kb and detected by at least eight markers were identified using the GENO-TYPING CONSOLE software (Affymetrix) and compared with variants reported in the Database of Genomic Variants.

Halgren et al.

Patient 2 was examined with Agilent Oligoarray 400K, patient 3 with Affymetrix 250K SNP array, patients 4 and 8 with Agilent 44K, patient 5 with Affymetrix 250K and Illumina Sentrix HumanHap300, patient 6 with Agilent Human Genome CGH Microarray 44 B, and patient 7 with Agilent Oligoarray 244K.

Results

All position coordinates given below are based on Human Feb. 2009 (GRCh37/19) assembly.

Clinical reports

Clinical data are provided in Table 1. Full clinical reports are provided in Appendix S1, Supporting information.

Patient 1

Patient 1 is an 8-year-old male. He was the first child of healthy unrelated parents. Routine second trimester ultrasound examination showed enlarged cerebral ventricles; amniocentesis was performed and a de novo balanced reciprocal translocation t(1;6)(p31;q25) was detected. The patient was born at term with a birth weight of 2450 g, birth length of 48 cm, and occipital frontal circumference (OFC) of 33 cm. He was hypotonic and had mild dysmorphic features. Developmental milestones were significantly delayed; he sat at the age of 2, walked at $2^{1/2}$, and spoke one to two words at the age of 3. Feeding problems and severe constipation were prominent from 6 months until 3 years of age. At the latest clinical examination, 8 years old, height was 120 cm, weight 21.8 kg, and OFC 54 cm. Dysmorphic features included a small triangular face, low hairline, micrognathia, small pointed chin, low-set large ears, broad nasal bridge and tip, concave curved thin vermilion of the upper lip, camptodactyly, and deep longitudinal plantar creases between first and second toes. He spoke only few words. Magnetic resonance imaging (MRI) showed complete ACC, and he was diagnosed with intellectual disability and autism according to the Global Assessment of Psychosocial Disability (GAPD) and the Autism Developmental Observation Schedule 1 (ADOS-1).

Patients 2-8

All seven patients had intellectual disability and speech impairment. Brain MRI was performed in four patients; partial ACC was detected in two of these, one had CC hypoplasia, while one patient had normal CC. Two of the patients had autism spectrum disorder (ASD) and autistic traits were found in another two. Hypotonia was reported in three patients, and four patients had feeding problems or failure to thrive in infancy.

Chromosome analysis

Chromosome analysis confirmed the karyotype 46,XY, t(1;6)(p31;q25) in patient 1 (Fig. 1a).

Next-generation paired-end sequencing

A total of 50,180,447 paired reads were generated in a single sequencing lane. Among these, 36,082,376 paired reads passed the chastity filter, 18,515,697 paired reads aligned uniquely, and 597,838 were chimeric pairs (end-reads mapping to different chromosomes). We removed non-clustering chimeric pairs leaving a total of 311 predicted breakpoints genome-wide that were visually filtered against known variants. The translocation breakpoints, resolved from three reads, were identified at 1q31.1 and 6q25.3 (Fig. 1b), leading to the refined karyotype 46, XY, t(1;6)(p31.1;q25.3). The breakpoint at 6q25.3 truncated intron 5 of ARID1B (RefSeq transcript NM_020732.3) while the chromosome 1 breakpoint affected no genes. Sanger sequencing identified the exact genomic positions of the breakpoints at chr1:73,895,566-73,895,579 and chr6:157,292,076-157,292,079. Four base pairs (TAGA) of unknown origin were inserted at the chromosome 1 breakpoint, while 23 base pairs (TCTGCAG AAAGTATAGGTCTGAT) were inserted at the chromosome 6 breakpoint; 22 of these (TCTGCAGAAAGT ATAGGTCTGA) match uniquely to a LINE sequence at chromosome 7 (Fig. 1c).

Quantitative polymerase chain reaction

Expression of *ARID1B* was observed in all analyzed subjects. The control samples were averaged, and the expressional levels in patient 1 using primers downstream of the translocation site and spanning the translocation were roughly half compared to the controls (Fig. 2). Expression data obtained with a primer set located upstream of the translocation showed that the expression in patient 1 was roughly twice compared to the controls.

Microarray analysis

No potentially pathogenic CNVs were detected in patient 1. In patient 2, a 0.2-Mb intragenic deletion in *ARID1B* was detected. Deletions in patients 3 and 4 only involved *ARID1B* while patients 5-8 all had larger deletions involving 5-73 RefSeq genes. Detailed information is provided in Fig. 3 and Table 1.

Discussion

Using NGS we showed that *ARID1B* at 6q25.3 was truncated in a patient carrying a *de novo* balanced translocation t(1;6)(p31.1;q25.3). The patient had ACC, intellectual disability, speech impairment, ASD, and mild dysmorphic features. To delineate the clinical features associated with haploinsufficiency of *ARID1B*, we compared the translocation patient to seven patients with overlapping interstitial *de novo* deletions.

We included all available patients in this study. Three patients had deletions that only affected *ARID1B*. Four patients had larger deletions encompassing 5–73 RefSeq genes, thus haploinsufficiency of other genes is

				Dotiont A	Dotiont 6	Dotiont 6	Dationt 7	Dotiont 0
		raugur 2		רמוופו וו 4				
Gender	Σ	Ŀ	Σ	Ш	Ц	ш	L	Ŀ
Age (years)	6	e	46	00	4	20	6	10
Birth (gestational week)	39	40	Ι	39	39	40	41	39
Birth length (cm)	48	49	Ι	44.5	Ι	50	46	45.5
Birth weight (g)	2450	3090	Ι	2390	2770	3050	2890	2640
OFC at birth (cm)	33	I	I	32	Ι	33	33.5	32.5
Clinical findings								
ID/DD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ACC	Yes	Not examined	Not examined	No	Not examined	Partial	Partial	Hypoplasia
ASD	Yes	Autistic traits	Autistic traits	Yes	Ι	Yes	No	No
Severe speech	Yes	Yes	Absent speech	Yes	Absent speech	Absent speech	Yes	Yes
impairment								
Seizures	No	No	Ι	No	No	Yes	Yes	Yes
Hypotonia	Yes	Yes	Ι	Yes	Yes	Yes	Yes	Yes
Low hairline	Yes	I	Ι	Yes	Ι	Yes	Yes	Yes
Low-set ears	Yes	Ι	Ι	Ι	Ι	Ι	Ι	Yes
Broad nasal tip	Yes	I	Ι	Ι	Ι	Yes	Ι	Yes
Thin vermilion of upper	Yes	Yes	No		I	Yes	I	I
lip								
Hypertrichosis	No	Yes	I		I	Ι	Yes	Yes
Pectus excavatum	No	Yes	I		I	Ι	I	Yes
Joint laxity	No	Yes	Yes	I	Ι	Ι	Ι	Yes
Vision	Hypermetropia	I	Myopia	Strabismus	I	Myopia, strabismus	Hypermetropia,	Cataracts
Growth							nystagmus	
Feeding problems in infancy/failure to thrive	Yes	Yes	I	I	Yes	Yes	Yes	I
Height (age)	117 cm (-1.8 SD) (7.5 years) 120 cm (-1.5 SD) (8 vears)	84 cm (-2.5 SD) (2 years 11 months)	1.59 m (–3.5 SD) (adult)	I	66.8 cm (–2 SD) (9 months)	152 cm (fifth centile) (18 years)	124 cm (–2 SD) (8 years 9 months)	112 cm (–4 SD) (9.5 years)
Weight (age)	20 kg (-1.6 SD) (7.5 years) 21 8 kn (8 vears)	12 kg (–1.2 SD) (2 years 11 months)	71 kg (adult)	I	5.6 kg (–2 SD) (9 months)	48 kg (10–25th centile) (18 years)	26.5 kg (–0.5 SD) (8 years 9 months)	19.5 kg (–2 SD) (9.5 years)
OFC (age)	54 cm (+1.4 SD) (7.5 years) 54 cm (8 years)	42.5 (-0.7 SD) (8 months)	56.5 cm (–0.75 SD) (adult)	I	43.9 cm (0 SD) (9 months)	54 cm (25–50th centile) (18 years)		48.5 cm (–2.5 SD) (9.5 years)

Brain malformations in patients with haploinsufficiency of ARID1B

Patient 1Patient 2Patient 3Patient 4Patient 5Patient 6Patient 7Patient 8Molecular characterization Karyotype $46, XY, 4(1;6)$ $46, XY, 4(1;6)$ $46, XY, 46, YY, 47, YY, $									
Molecular characterization Molecular characterization 46,XY, (11:6) 46,XX		Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8
Karyotype 46,XY,t(1;6) 46,XX	Molecular characteriz	ation							
Microarray platform Affymetrix Agilent Human Agilent Unan Agilent Uligoarray Agilent 44K NP 6.0 400K SNP array SNP array Agilent Vigoarray Agilent 44K SNP 6.0 400K SNP array SnP array Illumina Sentrix Genome CGH 244K Deletion (size) ^b No pathogenic 0.2 Mb 0.6 Mb 1.0 Mb 2.7 Mb 4.6 Mb 8.2 Mb 14.5 Mb Minimal deletions 157,215,659- 157,126,309- 156,423,608- 155,804,649- 152,788,282-157, 151,027,655-159, 153,185,096- Minimal deleted - 157,458,744 157,454,197 158,502,590 435,837 (152, 151,027,655-159, 153,073,486- foromosome 6 (157,210,495- (156,190,443- (155,797,565- 497,968-157, 019,422-159, (157,073,486- foromosome 157,465,1030 158,675) 158,672,590 158,672,590 158,517,307 966,910,1 167,754,366-	Karyotype	46,XY,t(1;6) (p31.1;q25.3)dn	46,XX	46,XY	46,XX	46,XX	46,XX	46,XX	46,XX
Deletion (size) ^b No pathogenic 0.2 Mb 0.6 Mb 1.0 Mb 2.7 Mb 4.6 Mb 8.2 Mb 14.5 Mb deletions 157,215,659- 157,126,309- 156,423,608- 155,804,649- 152,788,282-157, 151,027,655-159, 153,185,096- 167,727,387 Minimal deleted - 157,458,744 157,761,083 157,454,197 158,502,590 435,837 (152, 179,698 (151, 167,727,387 region on (157,210,495- (157,079,676- (156,190,443- (155,797,565- 497,968-157, 019,422-159, (153,073,486- foenomic position. 157,467.930 158,076,922) 158,517.307) 996,910) 187,660) 167,754,128)	Microarray platform	Affymetrix SNP 6.0	Agilent Oligoarray 400K	Affymetrix 250K SNP array	Agilent 44K	Affymetrix 250K; Illumina Sentrix HumanHap300	Agilent Human Genome CGH Microarray 44B	Agilent Oligoarray 244K	Agilent 44K
Minimal deleted 157,215,659- 157,126,309- 156,423,608- 155,804,649- 152,788,282-157, 151,027,655-159, 153,185,096- region on 157,458,744 157,761,083 157,454,197 158,502,590 435,837 (152, 179,698 (151, 167,727,387 region on (157,210,495- (157,079,676- (156,190,443- (155,797,565- 497,968-157, 019,422-159, (153,073,486- chromosome 6 (157,210,495- (157,079,676- (156,190,443- (155,797,565- 497,968-157, 019,422-159, (153,073,486- cenomic position. 157,467.930 158,076,922) 158,517.307) 996,910) 187,660) 167,754,128)	Deletion (size) ^b	No pathogenic deletions	0.2 Mb	0.6 Mb	1.0 Mb	2.7 Mb	4.6 Mb	8.2 Mb	14.5 Mb
hg19) ^c Affected gene(s) ^d <i>ARID1B</i> , <i>ARID1B</i> , <i>ARID1B</i> , <i>5</i> RefSeq genes 15 RefSeq genes 33 RefSeq genes 73 RefSeq genes incl. <i>ARID1B</i> incl. <i>ARID1B</i> , incl. <i>ARID1B</i> incl. <i>ARID1B</i> incl. <i>ARID1B</i> incl. <i>ARID1B</i> incl. <i>ARID1B</i> incl. <i>ARID1B</i> exons 1–7	Minimal deleted region on chromssome 6 (genomic position, hg19) ^c Affected gene(s) ^d	– ARID1B, intron 5	157,215,659- 157,458,744 (157,210,495- 157,467,930) <i>ARID1B</i> , exons 4–8	157,126,309- 157,761,083 (157,079,676- 157,806,675) <i>ARID1B</i> , exons 2–20	156,423,608- 157,454,197 (156,190,443- 158,076,922) <i>ARID1B</i> , exons 1–8	155,804,649- 158,502,590 (155,797,565- 158,517,307) 5 RefSeq genes incl. <i>ARID1B</i>	152,788,282-157, 435,837 (152, 497,968-157, 996,910) 15 RefSeq genes incl. <i>ARID1B</i> , exons 1-7	151,027,655-159, 179,698 (151, 019,422-159, 187,660) 33 RefSeq genes incl. <i>ARID1B</i>	153,185,096- 167,727,387 (153,073,486- 167,754,128) 73 RefSeq genes incl. <i>ARID1B</i>

ACC, Agenesis of corpus callosum; ASD, autism spectrum disorder; DD, developmental delay; F, female; ID, intellectual disability; M, male; OFC, occipital frontal circumference; SD, standard deviation.

^aWhen no information was available regarding a specific clinical feature the item was not scored. ^bAll deletions were *de novo*. ^cTheoretical maximal deleted region is given in brackets. ^dARID1B RefSeq transcript NM_020732.3.

Table 1. Continued



Fig. 1. Cytogenetic and molecular characterization of patient 1. (a) Partial karyotype, showing a *de novo* balanced reciprocal translocation involving chromosomes 1 and 6. Cytogenetic karyotype: 46,XY,t(1;6)(p31;q25)dn. Blue arrows indicate the cytogenetically determined breakpoints. (b) The translocation breakpoints were mapped using next-generation mate-pair sequencing. The chromosome 1 breakpoint mapped within a 3 kb non-genic region at 1p31.1 (shaded area) and the chromosome 6 breakpoint mapped within a 700 bp genomic region at 6q25.3 (shaded area), truncating *ARID1B*. The breakpoints were detected by three reads shown in green and blue (the colors indicate the strand orientation of the reads). (c) By Sanger sequencing, the exact genomic positions of the breakpoints were identified to chr1:73,895,566-73,895,579 and chr6:157,292,076-157,292,079 (hg19). Four base pairs (TAGA) of unknown origin were inserted at the (der)(1) breakpoint (shaded area) and 23 base pairs (TCTGCAGAAAGTATAGGTCTGAT) were inserted at the der(6) breakpoint (shaded area); 22 of these (TCTGCAGAAAGTATAGGTCTGA) match uniquely to a LINE sequence on chromosome 7.



Fig. 2. Expression pattern for *ARID1B* in patient 1 compared to five controls. Expressional levels in patient 1 using primers downstream of the translocation site were roughly half of that in the five controls. Expression levels in patient 1 using a primer set spanning the translocation likewise showed that expression was halved. Expression data obtained with a primer set located upstream of the translocation showed that the expression in patient 1 was roughly twice that of the average of the controls.

likely to impact the observed phenotypes. Brain MRI was not performed on three of the reported patients;

as this procedure would require general anesthesia it was decided against for ethical reasons. Despite these obvious limitations, overlapping clinical manifestations were present: all eight patients had intellectual disability, severe speech impairment, and various degrees of dysmorphic features. Callosal abnormalities were present in four of the five patients where brain imaging was performed. Three patients were diagnosed with ASD and another two showed autistic traits. This is in accordance with two recently published reports describing (i) a small de novo deletion within ARID1B in a patient with autism (15) and (ii) a patient with ACC, intellectual disability, speech impairment, and autism, in which a de novo translocation disrupted two genes: ARID1B and MRPP3 (16). Patient 4 had normal brain MRI; this is not surprising as ACC associated loci are known to exhibit reduced penetrance (8, 11, 17). Interestingly, that same patient had intellectual disability, speech impairment, and ASD, suggesting that these traits might not be associated with visible structural brain abnormalities.

(a) 155 Mb 165 Mb 160,Mb 6025.3 6027 6005 1 BAOK S chr1:73894514 c chr1:73897482 c236 Nord OPRM1 H TFB1M NOX3 MYCT1 ARIDIB H TULP4 H PACRO C6orf118 TTLL2 AKAP12 \$002 VIP OPRM1 H PDE10A CCR6 AKAP12 ARIDIB \$002 PACRG **ZBTB**2 MAS 41-8-1 DYNLT1 \$002 I PLG C60rf176 # EMND1 FBX05 CNKS C60rf35 | GPR31 OPRM1 EZR IGF2R C60rf176 | UNC936 ESR1 H OSTCL OPRM1 OPRM1 00285796 ESR1 SI C2263 H L0C4411771 UNC938 2066014 PRR18 RSPH LP MIRS Zp451B082 ESR1 OPRM1 SYNJ2 NTAP AGPAT4 OKI H SFT2D1 MLLT4 OPRM1 SYNJ2 NTAP LOC100289495 HGC6 TIAM2 SYNE1 OPRM1 SERAC1 WTAP 6820 BRP44L KIF25 SCAFS H GTF2H5 ACAT2 RPS6KA2 RPS6KA2 KIF25 QKI MTRF 1L TIAM2 TCP1 OKI SYTL3 MTRE 11 CLDN20 TCPI MIR191 RNASET2 RGS17 SLC22A2 060 OPRM1 SYTL3 SYTL3 LPALS MIR3 C60rf97 OPRM1 FGFR10P OPRM1 MIR3918 MAP3K4 FGFR10P OPRM1 TCP10 IPCEF 1 TROOP IPCEF TAGAP C6orf124 IPCEF1 TAGAP FNDC1 LOC1 00129518 SNOR828 MRPL18 L0C729683 SLC22A1 SLC22A1 (b) 6q25.3 chr1:73897482_c236 Nord Backx Genes RefSed ARID18 ARID18

Halgren et al.

Fig. 3. Detailed view of breakpoints and interstitial deletions affecting *ARID1B*. (a) Positions of interstitial *de novo* deletions affecting *ARID1B* in patients 2-8 (red bars). Green and blue 'chr1' reads illustrate the chromosome 6 breakpoint in patient 1. Black bars show the breakpoint previously published by Backx et al. and an intragenic deletion published by Nord et al. (b) *ARID1B* is disrupted in all eight patients. The figure was drawn according to the UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly.

As *ARID1B* is disrupted by the translocation in patient 1, the expression of the gene could be expected to be halved unless compensatory expression was done from the normal chromosome. The observed expression pattern for primer sets downstream of or spanning the translocation was in concordance with these expectations. As data sets for exons 5-7 and 7-8 exhibit virtually identical relative expression levels, it can be indirectly inferred that the downstream fragment carrying *ARID1B*, translocated onto the derivative

chromosome 1, is transcriptionally inactive as would be expected because this fragment carries no promoter region. The expressional pattern of exons located upstream of the translocation (exons 4-5) indicates that *ARID1B* is expressed at levels higher than for amplicons downstream of the translocation. This indicates that *ARID1B* is not only transcriptionally active on the normal chromosome but also from the fragment on der(6) which contains the intact promoter region. It is thus anticipated that the *ARID1B* fragment on der(1)

Brain malformations in patients with haploinsufficiency of ARID1B

is involved in the transcription of a chimeric mRNA consisting of the first five exons of ARID1B and an unknown part of chromosome 1. Interestingly these data are in concordance with previously published data from Backx et al. (16) who showed that a fusion transcript between ARID1B and MRPP3 was upregulated twofold in a patient with t(6;14) and a similar phenotype. The reason for the upregulation of ARID1B on der(6) can only be speculated, but a TargetscanS analysis (UCSC Genome Browser http://genome.ucsc.edu/) showed the presence of multiple putative miRNA regulatory sites in the 3' untranslated region of ARID1B. The lack of these miRNA regulatory sites on der(6) could easily be thought to loosen the expressional control of ARID1B potentially exerted by these putative regulatory sites, thus leading to increased expression from this allele. More work to confirm this theory, however, is needed.

Our findings emphasize that ARID1B is important for normal human brain development and function. ARID1B is a highly conserved gene which furthermore is associated with an evolutionary conserved stable gene desert, a hall mark of key developmental genes (18). It encodes a DNA-binding protein, ARID1B, that is part of the chromatin-remodeling complex SWI/SNF (19). Chromatin-remodeling complexes are involved in gene expression regulation. They act by altering the nucleosome structure which leads to changes in the chromatin structure that allows binding of transcriptional factors. Arid1b is expressed in the developing mouse brain (16) and studies of mouse embryonic stem cells have found Arid1b (BAF250b) to be particularly important in early development. Levels of BAF250b complexes were found to be high in undifferentiated mouse embryonic stem cells and lower during embryonic stem cell differentiation. Furthermore, BAF250b-deficient mouse embryonic stem cells were less capable of self-renewal and showed increased levels of differentiation (20, 21). Additional functional studies including a systematic search for ARID1B target genes may show how haploinsufficiency of ARID1B predispose to CC defects and to an array of cognitive defects, including severe speech defects.

Supporting Information

The following Supporting information is available for this article: Table S1. Primers used for Q-PCR examination of expressional levels of *ARID1B* in patient 1.

Appendix S1. Clinical reports of patients 2-8.

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

Please note: Wiley-Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

We thank the patients and families for participating. We thank Elisabeth Larsen, Theresa Wass, and Ingrid Kjær for technical assistance. Consultant pediatrician Karen Tilma Andersen is thanked for providing clinical data. This work was supported by the Danish National Research Foundation, the Lundbeck Foundation, and 'Cell Lines and DNA Bank of Rett syndrome, X mental retardation and other genetic diseases' (Medical Genetics-Siena) – Telethon Genetic Biobank Network (Project No. GTB07001C to AR).

References

- Paul LK, Brown WS, Adolphs R et al. Agenesis of the corpus callosum: genetic, developmental and functional aspects of connectivity. Nat Rev Neurosci 2007: 8 (4): 287–299.
- Wang LW, Huang CC, Yeh TF. Major brain lesions detected on sonographic screening of apparently normal term neonates. Neuroradiology 2004: 46 (5): 368–373.
- Bedeschi MF, Bonaglia MC, Grasso R et al. Agenesis of the corpus callosum: clinical and genetic study in 63 young patients. Pediatr Neurol 2006: 34 (3): 186–193.
- Glass HC, Shaw GM, Ma C, Sherr EH. Agenesis of the corpus callosum in California 1983–2003: a population-based study. Am J Med Genet A 2008: 146A (19): 2495–2500.
- Mangione R, Fries N, Godard P et al. Neurodevelopmental outcome following prenatal diagnosis of an isolated anomaly of the corpus callosum. Ultrasound Obstet Gynecol 2011: 37 (3): 290–295.
- Schell-Apacik CC, Wagner K, Bihler M et al. Agenesis and dysgenesis of the corpus callosum: clinical, genetic and neuroimaging findings in a series of 41 patients. Am J Med Genet A 2008: 146A (19): 2501–2511.
- Dobyns WB. Absence makes the search grow longer. Am J Hum Genet 1996: 58 (1): 7–16.
- O'Driscoll MC, Black GC, Clayton-Smith J, Sherr EH, Dobyns WB. Identification of genomic loci contributing to agenesis of the corpus callosum. Am J Med Genet A 2010: 152A (9): 2145–2159.
- Pirola B, Bortotto L, Giglio S et al. Agenesis of the corpus callosum with Probst bundles owing to haploinsufficiency for a gene in an 8 cM region of 6q25. J Med Genet 1998: 35 (12): 1031–1033.
- Rubtsov N, Senger G, Kuzcera H et al. Interstitial deletion of chromosome 6q: precise definition of the breakpoints by microdissection, DNA amplification, and reverse painting. Hum Genet 1996: 97 (6): 705–709.
- Nagamani SC, Erez A, Eng C et al. Interstitial deletion of 6q25. 2-q25.3: a novel microdeletion syndrome associated with microcephaly, developmental delay, dysmorphic features and hearing loss. Eur J Hum Genet 2009: 17 (5): 573–581.
- Firth HV, Richards SM, Bevan AP et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using ensemble resources. Am J Hum Genet 2009: 84 (4): 524–533.
- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memoryefficient alignment of short DNA sequences to the human genome. Genome Biol 2009: 10 (3): R25.
- Vandesompele J, De Preter K, Pattyn F et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002: 3 (7): RESEARCH0034.
- Nord AS, Roeb W, Dickel DE et al. Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur J Hum Genet 2011: 30: 1–5.
- Backx L, Seuntjens E, Devriendt K, Vermeesch J, Van Esch H. A balanced translocation t(6;14)(q25.3;q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum. Cytogenet Genome Res 2011: 132 (3): 135–143.
- van Bon BW, Koolen DA, Borgatti R et al. Clinical and molecular characteristics of 1qter microdeletion syndrome: delineating a critical region for corpus callosum agenesis/hypogenesis. J Med Genet 2008: 45 (6): 346–54.
- Ovcharenko I, Loots GG, Nobrega MA, Hardison RC, Miller W, Stubbs L. Evolution and functional classification of vertebrate gene deserts. Genome Res 2005: 15 (1): 137–145.
- Nagl NG Jr, Wang X, Patsialou A, Van Scoy M, Moran E. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. EMBO J 2007: 26 (3): 752–763.
- Yan Z, Wang Z, Sharova L et al. BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. Stem Cells 2008: 26 (5): 1155–1165.
- Li XS, Trojer P, Matsumura T, Treisman JE, Tanese N. Mammalian SWI/SNF-a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. Mol Cell Biol 2010: 30 (7): 1673–1688.