

RESEARCH ARTICLE

LRFN5 locus structure is associated with autism and influenced by the sex of the individual and locus conversions

Helle Lybæk¹ | Michael Robson² | Nicole de Leeuw³ | Jayne Y. Hehir-Kwa⁴ | Aaron Jeffries⁵ | Bjørn Ivar Haukanes¹ | Siren Berland¹ | Diederik de Bruijn³ | Stefan Mundlos² | Malte Spielmann⁶ | Gunnar Houge^{1,7,8} 

¹Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway

²Max Planck Institute for Molecular Genetics, Berlin, Germany

³Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

⁴Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands

⁵University of Exeter, Exeter, UK

⁶Institute for Human Genetics, UKSH, Kiel, Germany

⁷Institute of Clinical Medicine K2, Faculty of Medicine, University of Bergen, Bergen, Norway

⁸Honorary Chair of Evolution and Genomic Sciences, University of Manchester, Manchester, UK

Correspondence

Gunnar Houge, Department of Medical Genetics, Haukeland University Hospital, N-5021 Bergen, Norway.
Email: gunnar.douzgos.houge@helse-bergen.no

Funding information

Helse Vest, Grant/Award Number: 911459

Abstract

LRFN5 is a regulator of synaptic development and the only gene in a 5.4 Mb mammalian-specific conserved topologically associating domain (TAD); the *LRFN5* locus. An association between locus structural changes and developmental delay (DD) and/or autism was suggested by several cases in DECIPHER and own records. More significantly, we found that maternal inheritance of a specific *LRFN5* locus haplotype segregated with an identical type of autism in distantly related males. This autism-susceptibility haplotype had a specific TAD pattern. We also found a male/female quantitative difference in the amount histone-3-lysine-9-associated chromatin around the *LRFN5* gene itself ($p < 0.01$), possibly related to the male-restricted autism susceptibility. To better understand locus behavior, the prevalence of a 60 kb deletion polymorphism was investigated. Surprisingly, in three cohorts of individuals with DD ($n = 8757$), the number of deletion heterozygotes was 20%–26% lower than expected from Hardy–Weinberg equilibrium. This suggests allelic interaction, also because the conversions from heterozygosity to wild-type or deletion homozygosity were of equal magnitudes. Remarkably, in a control group of medical students ($n = 1416$), such conversions were three times more common ($p = 0.00001$), suggesting a regulatory role of this allelic interaction. Taken together, *LRFN5* regulation appears unusually complex, and *LRFN5* dysregulation could be an epigenetic cause of autism.

Lay Summary

LRFN5 is involved with communication between brain cells. The gene sits alone in a huge genomic niche, called the *LRFN5* locus, of complex structure and high mammalian conservation. We have found that a specific locus structure increases autism susceptibility in males, but we do not yet know how common this epigenetic cause of autism is. It is, however, a cause that potentially could explain why higher-functioning autism is more common in males than females.

KEYWORDS

allelic interaction, autism, chromatin structure, epigenetics, epigenomics, *LRFN5*, *SALM5*, TAD structure

INTRODUCTION

Autism spectrum disorders (ASD) are 3–4 times more frequent in males than in females (Newschaffer et al., 2007),

and this sex difference is even more prominent in the high-functioning ASD group (previously called Asperger syndrome) (Gillberg et al., 2006). The cause of this sex discrepancy is unknown and not X-linked (Zhao

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Autism Research* published by International Society for Autism Research and Wiley Periodicals LLC.

et al., 2007). Despite high heritability (Devlin & Scherer, 2012; Lauritsen et al., 2005), the genetic causes of mild ASD are mostly unknown in the absence of pathogenic genomic copy number gains or losses (Sebat et al., 2007; Weiss et al., 2008). Genome-wide association studies (GWAS) have not identified major autism loci, but have linked ASD, intellectual disability (ID) and schizophrenia to an overlapping set of low-risk associated variants. The *LRFN5* locus, this article's target of investigation, was only listed in a family-based GWAS study from 2009 (Wang et al., 2009), a finding not replicated in more recent population-based or meta-GWAS analyses (Autism Spectrum Disorders Working Group of The Psychiatric Genomics, 2017; Grove et al., 2019). There are, however, other results that link the *LRFN5* locus to autism. A copy number variation (CNV) study of an ASD cohort found that nearly 1% of all ASD-linked rare CNVs were in the *LRFN5* locus and none were reported de novo (Pinto et al., 2010). Even more relevant is a homozygous-haplotype-sharing-in-autism-study that identified both *LRFN5* and the flanking gene *FBXO33* as ASD candidate genes in at least two population clusters (Casey et al., 2012).

LRFN5, also known as SALM5 (synaptic adhesion-like molecule-5) in mice, belongs to a family of five small transmembrane protein genes involved in synaptic development, organization and plasticity (Choi et al., 2016; Goto-Ito et al., 2018; Ko et al., 2006; Ko & Kim, 2007; Morimura et al., 2006; Seabold et al., 2008). *LRFN5* acts as a dimer, and induces presynaptic differentiation through binding to the LAR family of receptor tyrosine phosphatases (Lin et al., 2018). *LRFN5* is the only family member located in the middle of a gene desert that can also form a large topologically-associating domain (TAD) of 5.4 Mb. Such a domain constitutes a self-interacting genomic region, isolated from neighboring TADs by isolator sequences associated with CCCTC-binding factor sites, called CTCF-sites. A TAD can be regarded as a gene-regulatory unit, and within a TAD a gene can interact with one or more enhancers that regulate gene expression, sometimes in a time- and tissue-specific fashion. The single-gene-in-a-TAD structure is not unique, but this TAD is unusually large, the average TAD size being about 0.9 Mb (Yu & Ren, 2017). This suggests complex gene regulation at the structural level. It is also the only *LRFN*-gene with a large 5'UTR (of 1.9 kb), encoded by the first 2 (of 6) exons. This 5'UTR can potentially form a complex hairpin with an estimated ΔG of -760 kcal/mol (<http://www.unafold.org/>). This suggests gene regulation also at the translational level.

Inspired by rare cytogenetic findings in individuals with translocations or copy number aberrations affecting the *LRFN5* locus, one with higher functioning autism (Cappuccio et al., 2019), and one with severe ID with good motor function but no social communication (de Bruijn et al., 2010), in-depth investigation of locus structure was done to look for autism-associated changes,

and to explain sex-biased inheritance of autism susceptibility in distantly related males. We found that mild structural differences in this region are related to the sex of the individual and associated with autism in males, and that locus homogenization by frequent allelic conversions could both be a mechanism of structure maintenance and of *LRFN5* expression regulation. Complex locus regulation could possibly be related to fine-tuning of gene expression, for example, through establishment of monoallelic expression.

MATERIALS AND METHODS

Ethical board approvals and consent for publication

The methylation study on the anonymized ASD cohort was approved by the Ethical Review Board of Northern Norway, REK-Nord# 2013/965. The study on sex-dependent differences in the epigenetic profile of the *LRFN5* locus was approved by the Ethical Review Board of Western Norway, REK-Vest# 2016/25. All involved members of the two families with two autistic brothers, including the brothers themselves, have been counseled about these results and have given consent to publication, recorded in the hospital's journal. Other individuals participating in this study were anonymized.

Patients and DNA samples

DNA samples were obtained from the diagnostic genetic service in Bergen (Norway) and in Nijmegen (the Netherlands). Patients were recruited from the clinical genetic service in Bergen (Norway). Consent was obtained after genetic counseling of the two families having boys with autism. The senior author (Gunnar Houge) met these four males with autism 10–20 years ago and again recently when they were informed about the results presented here and our plans to publish. They all have learning problems of similar degree but not always similar type, and the same difficulties with social interaction, finding communication through the computer easier than to face-to-face. Despite finishing 9 years of elementary school and 1–3 years of high school with results in the average range, they are unable to hold regular jobs or participate in team sports (like football) because of their social handicap. Some has had education beyond high-school and also work training, so far without getting a job. The learning problems could affect both mathematics and languages, but to variable degrees. For some, a digital clock display is easier to understand than a classical clockface, and despite knowing the alphabet, the order of the letters represents a challenge. Their range of daily-life problems is strikingly similar, and they all give a similar impression when you talk with them.

Chromatin immunoprecipitation-on-chip experiments

Chromatin immunoprecipitation (chIP) was done on primary fibroblast cultures obtained from skin biopsies from patients, family members and controls. Fibroblasts were frozen in several batches after primary culture, and a single batch was thawed for an experiment. This was done to avoid major differences in number of passages (cell culture doubling times) between experiments. For each sample, a total of 4×10^7 fibroblast cells were cultured to ~90% confluence. Cells were cross-linked for 15 min with 1% formaldehyde at 37°C before 125 mM glycine was added, followed by incubation for 5 min at room temperature (RT). Next, cells were washed three times with ice-cold phosphate-buffered saline (PBS) before harvesting by scraping, and cell pellet were collected by centrifugation. Chromatin from cell pellet was isolated as follows: Incubation for 10 min on ice in 4 ml cell lysis buffer (5 mM PIPES pH 8; 85 mM KCl; 0.5% NP-40; 1× protease inhibitors), centrifugation and incubation of cell pellet for 30 min on ice in another lysis buffer (50 mM Tris pH 8; 10 mM EDTA; 1% SDS; 1× protease inhibitors). Chromatin was sheared on a S220 ultra-sonicator (Covaris, MA, USA) according to the protocol “Chromatin shearing with SDS detergent buffers” from the manufacturer. The level of fragmentation was examined on a 1% agarose gel, and shearing was continued until the major part of the DNA fragments had a size of ~600 bp. The cell solution was centrifuged, and the lysate was frozen in batches of 100 μ L at -80°C for later use, except for 50 μ L lysate which was used for determination of the chromatin concentration (=input DNA) by standard DNA isolation procedures.

chIP was done as follows: Per chIP, 50 μ L chromatin (on average 15 μ g), 25 μ L protein A/G PLUS beads (Santa Cruz Biotechnology, TX; USA), and 425 μ L incubation buffer (0.2% SDS, 1% Triton, 150 mM NaCl, 2 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.5, 1× protease inhibitors) were incubated with rotation for 1 h at 37°C. Protein-DNA complexes were immunoprecipitated with a chIP-validated antibody at 4°C. The antibodies used were against histone H3 (histone H3, #4620; Cell signaling, MA, USA), histone H3 trimethylated at lysine 4 (H3K4me3, ab8580; Abcam, Cambridge, UK), histone H3 trimethylated at lysine 27 (H3K27me3, ab6002; Abcam), histone H3 trimethylated at lysine 9 (H3K9me3, ab8898; Abcam), histone H3 acetylated at lysine 9 (H3K9ac, ab4441; Abcam), and rabbit IgG (#2729; Cell signaling). The amount of antibody added was according to manufacturers' recommendations. Next day, 25 μ L protein A/G PLUS beads were added to each chIP and incubated for 2 h at 4°C. Subsequently, the protein-DNA complexes underwent a series of washes: 2× (0.1% SDS, 0.1% DOC, 1% triton, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.5), 1× (0.1% SDS, 0.1% DOC, 1%

triton, 500 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.5), 1× (0.25% LiCl, 0.5% DOC, 0.5% NP-40, 1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.5), and 2× (1 mM EDTA, 0.5 mM EGTA, 10 mM Tris pH 8.5). The formaldehyde-induced crosslinks were reversed by incubation for 30 min at RT with SDS-solution (0.1 M NaHCO₃ + 1% SDS), and genomic DNA was recovered by standard DNA isolation procedures. Validation of the chIPs were done by PCR on non-amplified chIP samples with primers recognizing the *FBXO33* and *LRFN5* genes. In the PCR, rabbit IgG-chIP was used as a preimmune control and 2% of input DNA was used as a control of the immunoprecipitation efficiency.

For the chIP-on-chip analysis, 60% of each of two replicate chIP samples were concentrated using microcon YM-30 spin columns (Merck Millipore, MA, USA) and amplified using the GenomePlex Complete Whole Genome Amplification kit (Sigma-Aldrich, MO, USA). The chIP-on-chip hybridizations were done on a custom designed high resolution NimbleGen 3× 720 K array (Roche NimbleGen, WI, USA). The array probes (50-mers, positions according to Chr37/hg19) covered unique regions of chromosome 2 (168,500–178,500 Mb), chromosome 13 (94,000–113,000 Mb), chromosome 14 (16,475–70,975 Mb), and chromosome 17 (41,370–81,195 Mb) uniformly, with a median probe spacing of ~150 bp. DNA labelling, array hybridization, post-hybridization washes and scanning were performed according to the manufacturer's protocol: “NimbleGen Arrays User's Guide, ChIP-chip Array”, v6.2 (Roche NimbleGen). In short, the chIP and input DNA (DNA from nonprecipitated chromatin) samples were labeled with Cy5- and Cy3-conjugated random nonamers, respectively. The labeled samples were purified, combined, denatured and hybridized to the array for 16 h at 42°C. After stringent washing, the array was scanned using an Axon 4200AL Scanner (Molecular Devices, CA, USA) at 5- μ m resolution. The acquired images were analyzed by DEVA v1.2 software (Roche NimbleGen) creating pair reports, including raw intensities for each probe and per image. From these data, ratio files were generated. For data visualization, the average ratios of two replicate experiments were binned per kb, each adjusted for the number of probes per bin. These data were transferred to in Excel spreadsheets for further calculations and generation of plain text files in .bedgraph format, and then the data was plotted against chromosomal position using the UCSC browser's custom track option.

Capture HiC-based *LRFN5* locus TAD-structure determination

LRFN5 locus selection

The capture Hi-C (CHiC) SureSelect library was designed over the genomic interval (chr14:539,000,000-47,000,000,

GRCh37) using the SureDesign tool from Agilent (Agilent Technologies, Santa Clara, CA). The coverage was 70.5% by 159,698 probes of total size 4668 Mb and 5× tiling density.

Fixation of fibroblast nuclei

Capture HiC experiments were performed on dermal fibroblasts from a family trio (parents and child with ASD) and a control female of the same age as the mother. Trypsinized fibroblasts were washed in PBS and then transferred to a 50-ml Falcon tube and complemented with 10% FCS/PBS. 37% formaldehyde was added to a final concentration of 2% and cells were fixed for 10 min at RT. Crosslinking was quenched by adding glycine (final concentration; 125 mM). Fixed cells were washed twice with cold PBS and lysed using fresh lysis buffer (10 mM Tris, pH 7.5, 10 mM NaCl, 5 mM MgCl₂, 0.1 mM EGTA with protease inhibitor) to isolate nuclei. Cell lysis was assessed microscopically after 10-min incubation in ice. Nuclei were centrifuged for 5 min at 480 g, washed once with PBS and snap frozen in liquid N₂.

Chromosome conformation capture library preparation and sequencing

The 3C libraries were prepared from fixed nuclei as described previously (Kragesteen et al., 2018). Briefly, lysis buffer was removed by centrifugation at 400 g for 5 min at 4°C, followed by supernatant aspiration, snap-freezing, and pellet storage at -80°C. Later, nuclei pellets were thawed on ice, resuspended in 520 µL 1× DpnII buffer, and then incubated with 7.4 µL 20% SDS shaking at 900 rpm. at 37°C for 1 h. Next, 75 µL 20% Triton X-100 was added and the pellet was left shaking at 900 rpm. at 37°C for 1 h. A 15-µl aliquot was taken as a control for undigested chromatin (stored at -20°C). The chromatin was digested using 40 µL 10 U/µl DpnII buffer shaking at 900 rpm. at 37°C for 6 h; 40 µL of DpnII was added and samples were incubated overnight, shaking at 900 rpm. at 37°C. On day three, 20 µL DpnII buffer was added to the samples followed by shaking for 5 more hours at 900 rpm. at 37°C. DpnII subsequently was inactivated at 65°C for 25 min and a 50-µl aliquot was taken to test digestion efficiency (stored at -20°C). Next, digested chromatin was diluted in 5.1 ml H₂O, 700 µL 10 × ligation buffer (Thermo Fisher Scientific), 5 µL 30 U/µl T4 DNA ligase and incubated at 16°C for 4 h while rotating. Ligated samples were incubated for a further 30 min at RT. Chimeric chromatin products and test aliquots were de-crosslinked overnight by adding 30 µL and 5 µL proteinase K, respectively, and incubated at 65°C overnight. On the fourth day, 30 µL or 5 µL of 10 mg/ml RNase was added to the samples and aliquots,

respectively, and incubated for 45 min at 37°C. Next, chromatin was precipitated by adding 1 volume phenol-chloroform to the samples and aliquots, vigorously shaking them, followed by centrifugation at 4000 rpm. at RT for 15 min. To precipitate aliquoted chromatin, 1 volume 100% ethanol and 0.1 volume 3 M NaAc, pH 5.6 was added and the aliquots placed at -80°C for 30 min. DNA was then precipitated by centrifugation at 5000 rpm. For 45 min at 4°C followed by washing with 70% ethanol, and resuspension in 20 µL with 10 mM Tris-HCl, pH 7.5. To precipitate samples, extracted sample aqueous phases were mixed with 7 ml H₂O, 1 ml 3 M NaAc, pH 5.6, and 35 ml 100% ethanol. Following incubation at -20°C for at least 3 h, precipitated chromatin was isolated by centrifugation at 5000 rpm. For 45 min at 4°C. The chromatin pellet was washed with 70% ethanol and further centrifuged at 5000 rpm. For 15 min at 4°C. Finally, 3C library chromatin pellets were dried at RT and resuspended in 10 mM Tris-HCl, pH 7.5. To check the 3C library, 600 ng were loaded on a 1% gel together with the undigested and digested aliquots. The 3C library was then sheared using a Covaris sonicator (duty cycle: 10%; intensity: 5; cycles per burst: 200; time: 6 cycles of 60 s each; set mode: frequency sweeping; temperature: 4-7°C). Adaptors were added to the sheared DNA and amplified according to the manufacturer's instructions for Illumina sequencing (Agilent). The library was hybridized to the custom designed SureSelect beads and indexed for sequencing (150 bp paired-end) following the manufacturer's instructions (Agilent).

Capture-HiC analysis

Raw reads were preprocessed with cutadapt v1.15 to trim potential low-quality bases (-q 20 -m 25) and any remaining sequencing adapters (-a and -A option with Illumina TruSeq adapter sequences according to the cutadapt documentation) at the 3' ends of the reads. Mapping, filtering and deduplication of the short reads were performed with the HiCUP pipeline v0.7.04 (<https://doi.org/10.12688/2Ff1000research.7334.1>) (no size selection, Nofill: 1, Format: Sanger). The pipeline employed Bowtie2 v2.2.6 (<https://doi.org/10.1038/2Fnmeth.1923>) for mapping short reads to the hg19 human reference genomes. Juicer tools 0.7.5 (<https://doi.org/10.1016/2Fj.cels.2016.07.002>) was used to generate binned-contact maps from valid and unique read pairs with MAPQ ≥ 30 and to normalize contact maps by Knight and Ruiz (KR) matrix balancing (<https://doi.org/10.1016/2Fj.cell.2014.11.021> <https://doi.org/10.1093/2Fmanus/2Fdrs019>). For the generation of capture-HiC contact maps, only reads pairs mapping to the enriched genomic region were considered and shifted by the offset of the enriched genomic region. For the import with Juicer tools, we used a custom chrom.sizes file containing only the size of the enriched part of the genome. Afterward,

KR-normalized maps were exported at 10-kb resolution, and coordinates were shifted back to their original values.

DNA methylation studies

CpG methylation of the *LRFN5* locus was initially interrogated by the Illumina 450 K DNA methylation array (Infinum HumanMethylation450 BeadChip), following manufacturer's protocol. No methylation differences were found between male and female controls. Specific positions in the *LRFN5* locus (namely the region corresponding to the peak of the locus-TAD and the *LRFN5* promoter) were analyzed by both DNA sequencing after bisulfite treatment, and custom designed MLPA tests (using the P-300 kit from MRC-Holland, and following the manufacturer's protocol). Still, no sex

differences were found. All these methylation studies were done on leukocyte DNA from peripheral blood samples.

SNP array analysis of a deletion polymorphism upstream of *LRFN5*

The allele frequencies of a common 60 kb deletion, chr14 (GRCh37):g.41609383–41,669,664, detected by at least 12 oligonucleotides on the Affymetrix 6.0 SNP array and by at least 18 oligonucleotides on the Affymetrix CytoScan SNP array (Affymetrix, ThermoFischer Scientific, USA), were determined in four cohorts of individuals: three ascertained due to developmental disorders, and one group of mostly medical students (Table 2). All cohorts were anonymized. Three of the cohorts (patient cohort I and III and the student cohort) had their

TABLE 1 *LRFN5* locus haplotypes in affected individuals and their families, all having overlapping deletions on different haplotypes (A–E) just upstream of the *LRFN5* promoter

Fam #	Individual	DD/ASD	Deletionhaplotype ^a
1	Father	—	—
1	Mother	—	A
1	Son ^b	—	A
1	Son	ASD	A
1	Son	ASD	A
1	Son	—	—
1	Daughter	—	—
1	Daughter	—	—
2	Mat grandfather	—	—
2	Mat grandmother	—	A
2	Mat uncle	—	A
2	Father	—	—
2	Mother	—	A
2	Son	ASD	A
2	Son	ASD	A
2	Son	—	—
2	Daughter	—	—
3	Father	—	B
3	Mother	—	—
3	Son	DD	B
4	Father	—	—
4	Mother	—	C
4	Son	DD/ASD	C
5	Father	—	—
5	Mother	—	D
5	Daughter	Rett-like	D
6	Adult male	Mild ID/ schizophrenia	E

^aDeletion sizes (kb, hg19): A: 41846–42,020, B: 41846–42,023, C: 41867–41,979, D: 41822–41,933, E: 41836–42,004. The haplo-types were also determined based on *LRFN5* locus SNP pattern.

^bThis individual was a SRY-negative 46,XX-DSD male.

TABLE 2 Allelic distribution of a common 60 kb deletion^a (41608–41,657 kb based on SNP-array data) in three patient and one student cohort

	<i>n</i>	wt/ wt	wt/ del	del/ del	Deletion MAF	wt/wt	wt/ del	del/ del	Δ	% loss of wt/del to wt/wt and del/del
Patient cohort I (NO)	850	649	168	33	0.138	0.764	0.198	0.039		10% + 10%
<i>Expected if HWE</i>						0.743	0.238	0.019	2.1	
Patient cohort II (NO)	4843	3582	1025	236	0.155	0.740	0.212	0.049		12% + 12%
<i>Expected if HWE</i>						0.714	0.262	0.024	2.0	
Patient cohort III (NL)	3064	2600	389	75	0.088	0.849	0.127	0.024		13% + 13%
<i>Expected if HWE</i>						0.832	0.161	0.008	3.0	
Student cohort (NL)	1416	1247	114	55	0.079	0.881	0.081	0.039		40% + 40%
<i>Expected if HWE</i>						0.848	0.146	0.006	6.5	

Note: The Δ -column shows the fold difference of allele frequencies between observed del/del and expected del/del if Hardy–Weinberg equilibrium.

^aSize based on gnomAD v3.1.1 whole-genome sequencing data is 41,609,383–41,669,664 (60.3 kb) with MAF 0.139 (2990/21518 alleles).

genomic copy number status determined by the Affymetrix 6.0 array, while patient cohort II was tested by the Affymetrix CytoScan array.

Statistical analysis

Basic and on-line statistical tools were used to investigate the statistical significance of the following findings: the segregation of the A-haplotype in the autism family: Fisher Exact Test ($p = 0.0476$); the female–male comparison in the table of Figure 3: Student two-sided t test ($p = 0.005$); the difference between Hardy–Weinberg equilibrium and observed allelic distribution in the patient cohorts in Table 2: Chi-square test ($p < 0.00001$); and the difference between the allelic distribution in the patient cohorts and the student cohort in Table 2: Chi-square test ($p < 0.00001$).

RESULTS

Our attention was drawn to the *LRFN5* locus in 2007 when a girl with ID that gave no social contact was found to have two de novo chromosome changes: a small 2q31.1 deletion and a balanced 14;21-translocation (de Bruijn et al., 2010). Later we compared the chip-on-chip results from this patient's skin fibroblasts with the results from a nonautistic boy with another *LRFN5* locus translocation: a de novo t(6;14)(q26;q21.1) causing Coffin–Sirin syndrome because *ARID1B* was disrupted on chromosome 6 (Figure S1). A clear difference in the *LRFN5* locus chromatin profiles was only seen in the autistic girl.

Of more general relevance for *LRFN5* function is the number of individuals with DD/ASD and copy number variants (CNVs) in the *LRFN5* locus in the DECIPHER database ($n = 30$) and our own records ($n = 3$) (Figures S2 and S3). Only one small deletion

containing *LRFN5* itself was registered as de novo. Because many CNVs are inherited from a seemingly unaffected parent, most have been regarded as non-pathogenic. Given the results below, this assumption could be false.

Autism in males segregated with a specific *LRFN5* locus haplotype when maternally inherited

High-functioning ASD was diagnosed in two pairs of brothers from two families from the same geographical region (Table 1; fam #1 and fam #2). These four males have been seen by the senior author both as children and adults, and their autism problems are strikingly similar (for details, see Patients and DNA samples). They all had a 172 kb deletion just upstream of the *LRFN5* promoter and the same locus haplotype determined by SNP-array analysis, called the A-haplotype in Table 1. The A-haplotype was inherited from their normal mothers. The families were too distant to know about any relatedness. The A-haplotype was also found in a normal maternal uncle, inherited from his mother. Of the nine individuals sharing the A-haplotype in these families, 4/5 males had ASD, and 4/4 *SRY*-negatives (three females and one 46,XX-male) were normal. This suggests that the A-haplotype increased ASD susceptibility in males in this family (Fisher Exact Test $p = 0.0476$). We also found overlapping deletions on different haplotypes in four other individuals or families, ascertained by diagnostic SNP-array copy-number testing because of developmental delay, ID, ASD or schizophrenia (Table 1). This could indicate that the A-haplotype is linked to male ASD-susceptibility, not the 172 kb deletion per se. An effect of the deletion is not excluded, but then it must be haplotype dependent. The 172 kb deletion does not contain enhancer-like chromatin profiles or lncRNAs.

The structure of the *LRFN5* locus is influenced by the sex of the individual

To explore if the A-haplotype influenced TAD-structure of the *LRFN5* locus, we performed a capture HiC-experiment on skin fibroblasts from a family trio (boy with ASD and his unaffected parents), with a normal unrelated female as control (Figure 1). Three patterns emerged: In the autistic boy and his mother, both sharing the A-haplotype, the whole locus 5.4 Mb mega-TAD had three sub-TADs with boundaries at the *LRFN5* promoter (arrow A in Figure 1) and the middle of the *LRFN5*-downstream gene desert (arrow B in Figure 1). The A-haplotype-associated 172 kb deletion is marked with a

star in Figure 1, and it looks like a small red diamond. In the father, only two TADs could be discerned inside the mega-TAD. No distinct “B-junction” could be seen. In the control female, only the mega-TAD was distinct; an “A-junction” was diffuse if at all present (Figure 1).

These subtle capture-HiC differences point to variation in TAD structure and that the autistic boy inherited his mother’s structure. Possibly, this specific structure (with a “B-junction”) could be A-haplotype dependent and linked to autism in this family. We also noted that the father and the control female had TAD differences (Figure 1).

To find out if these TAD differences could be associated with differences in chromatin profiles, chIP-on-chip-

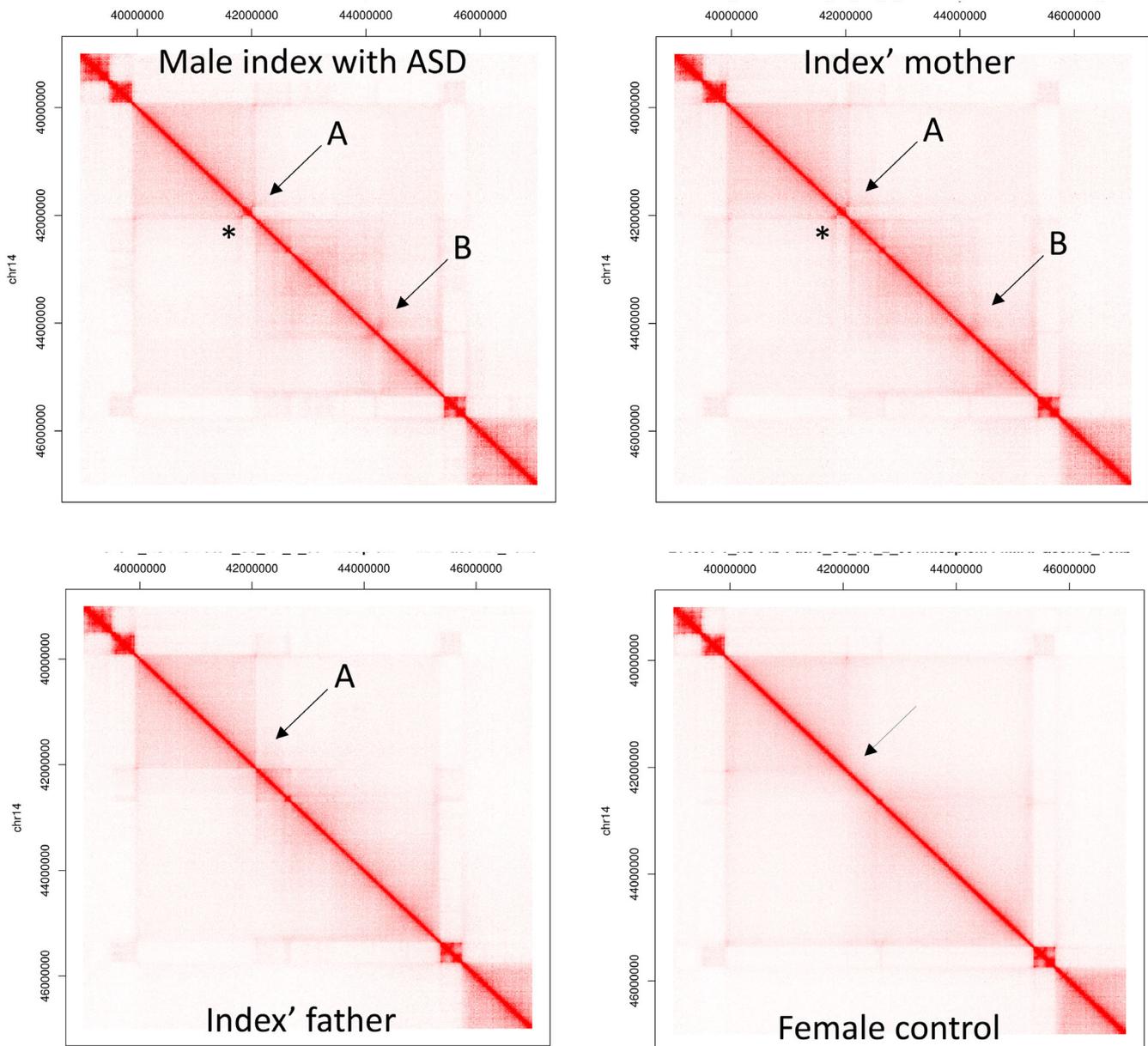


FIGURE 1 Capture-HiC results of a family trio (index male with ASD, his mother and father) and a control female (of the same age as index’ mother). Arrows indicate TAD junctions (A and B). The small diamond (marked with an asterisk) just centromeric to TAD junction (A) (corresponds to the *LRFN5* promoter) marks the 172 kb familial deletion

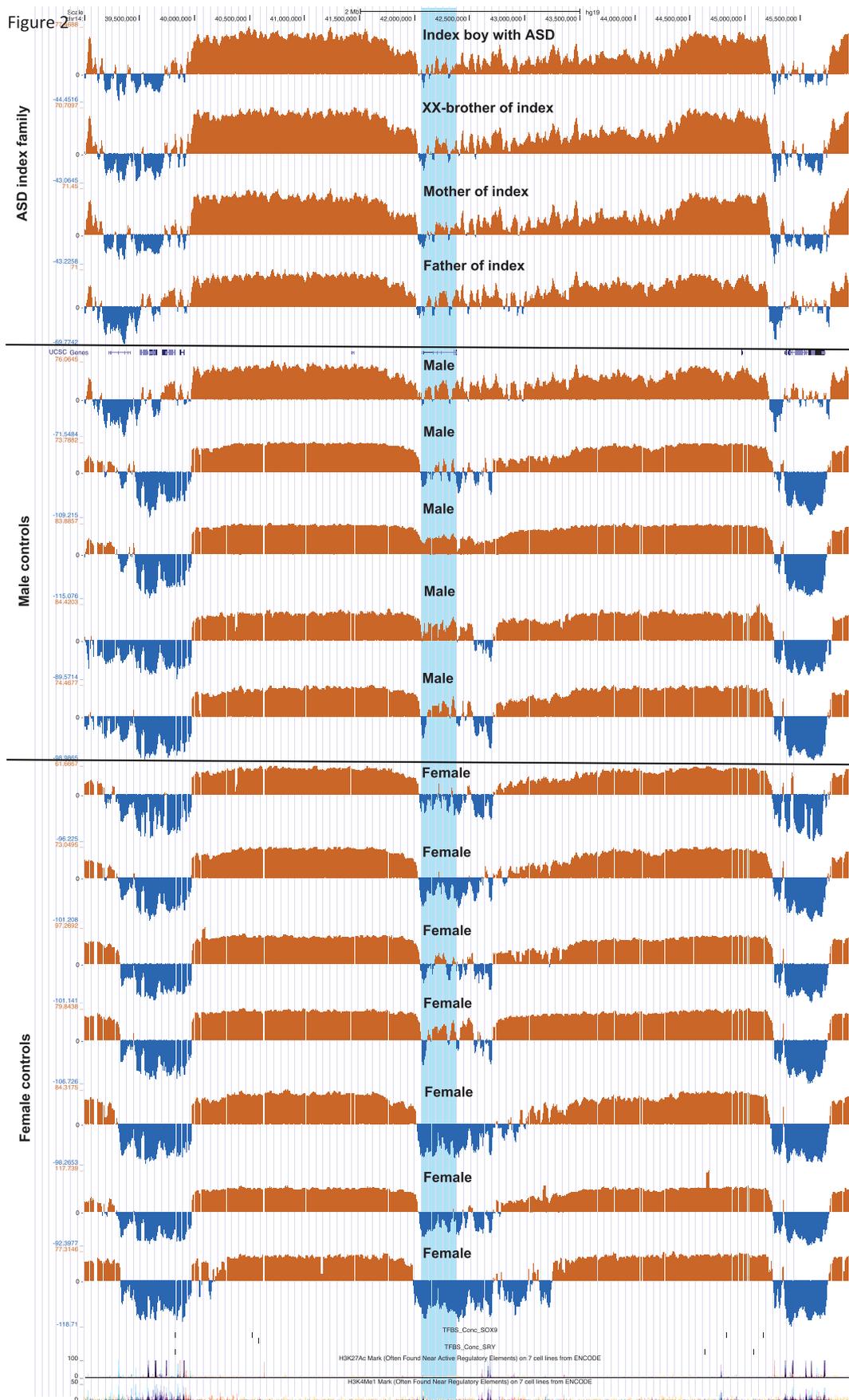


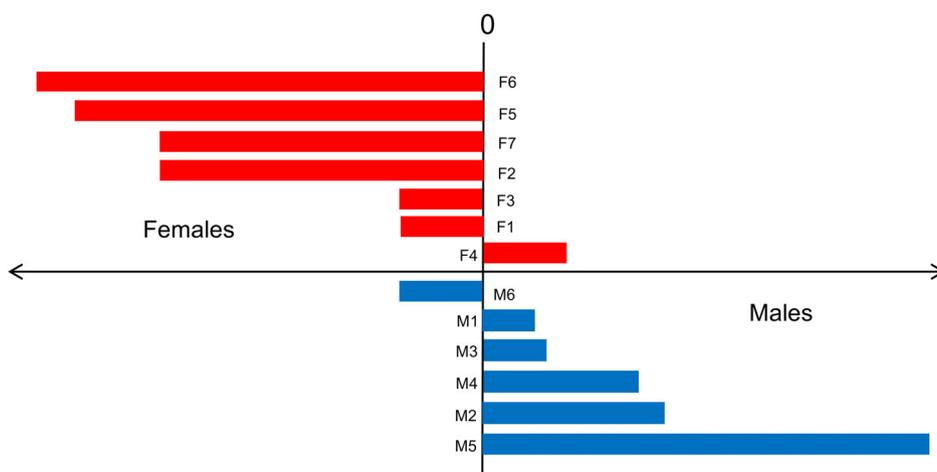
FIGURE 2 Top: *LRFN5* locus H3K9me3 chIP-on-chip profiles from index with ASD, his 46,XX-DSD nonautistic brother, his normal mother and his normal father (top four lanes). Middle: Profiles from five control males. Bottom: Profiles from seven control females. Note the profile variability around the *LRFN5* gene in contrast to the flanking H3K9me3-enriched domains

based chromatin quantification of H3K4me3, H3K27me3 and H3K9me3-associated chromatin was done in skin fibroblast cultures from four family members, five control males and seven control females. A sex difference was only found for the H3K9me3 profiles, that is, heterochromatin protein-1 (HP1)-associated constitutive heterochromatin, and only corresponding to the *LRFN5* gene itself and the region downstream to the H3K4me3/H3K27me3 signals marking the peak of the mega-TAD (Figure 2). In this area, males had more heterochromatin than females. This difference was significant, but individual differences in the degree of this effect should be noted (Figure 2). In fact, the groups are overlapping with one female (F4) having an average H3K9me3 profile value similar to two males (M1 and M3), and one male (M6) having an average H3K9me3 profile value similar to two females (F1 and F3) (Figure 3). In contrast, the sex difference was zero in a region 1 Mb upstream (Figure 3).

The four family members in Figure 2, including the mother and the *SRY*-negative brother of the index boy that also shared the A-haplotype (Table 1), all had a “male-type” of chromatin pattern with a stronger H3K9me3 signal around *LRFN5*. Hypothetically, this pattern could be *SRY*-dependent, but not if one inherits the A-haplotype. The control male and female with the

largest H3K9me3-associated chromatin differences are also shown at the bottom of the locus overview figure (Figure 4), where their H3K9me3-chromatin profiles are aligned to the ASD-associated A-haplotype TAD structure, CTCF-binding sites, and open-chromatin neuronal single-cell ATAC sites.

We also wanted to know if we could test large numbers of individuals for sex-related or autism-related differences in this region, and for practical reasons we then needed a test that could be done on blood leukocyte DNA, for example, a CpG methylation test. Accordingly, we performed a pilot MLPA-based study on anonymized blood leukocyte DNA to look for methylation differences of a CpG located at the conserved CTCF binding site in the *LRFN5* promoter (chr14:42,069,895-42,069,949, hg19). An average methylation degree of 0.19 (CI 0.13–0.25) in control males ($n = 16$) and control females ($n = 6$) were found with no sex difference, and the same average level (0.20) was found in patients investigated because of higher-functioning autism ($n = 14$). In a separate experiment on blood-DNA, using the Illumina 450 K BeadChip, we found an average degree of CpG methylation of 0.14 at the *LRFN5* promoter (19 CpGs were interrogated) and 0.52 in the *LRFN5* gene itself (8 CpGs were interrogated) in 10 other control individuals, still without sex



Display of the individual *LRFN5* chromatin data from males (M1–M6) and females (F1–F7) based on the table below

H3K9me3 chromatin scores (relative to genomic average) in fibroblasts from control males (M's) and females (F's)

Genomic position	M1	M2	M3	M4	M5	M6	F1	F2	F3	F4	F5	F6	F7	Average males	Average females	Diff
<i>LRFN5</i> (42050-42400 kb)	3	11	4	9	26	-5	-5	-19	-5	5	-24	-26	-19	8 (0 - 16)	-13 (-22 - -4)	21*
Reference pos. (40550-40900 kb)	54	57	51	38	68	60	46	58	60	61	43	60	55	55 (47 - 63)	55 (49 - 61)	0

The chromatin scores are based on 673-1907 measurement points (= one oligo on the array), on average around 700 measurements per position. The reference position was chosen to be in the middle of the heterochromatin stretch upstream of *LRFN5*. 95% confidence intervals are shown in parentheses. *p = 0.005 (Student two-sided t-test).

FIGURE 3 Quantification of data from Figure 2: *LRFN5* locus H3K9me3 chromatin levels (relative to genomic average) in fibroblasts from control males (M1–M6) and females (F1–F7). The investigated positions are indicated in the table below, and the bar diagram above illustrates the individual male/female distribution. Note that the groups are overlapping

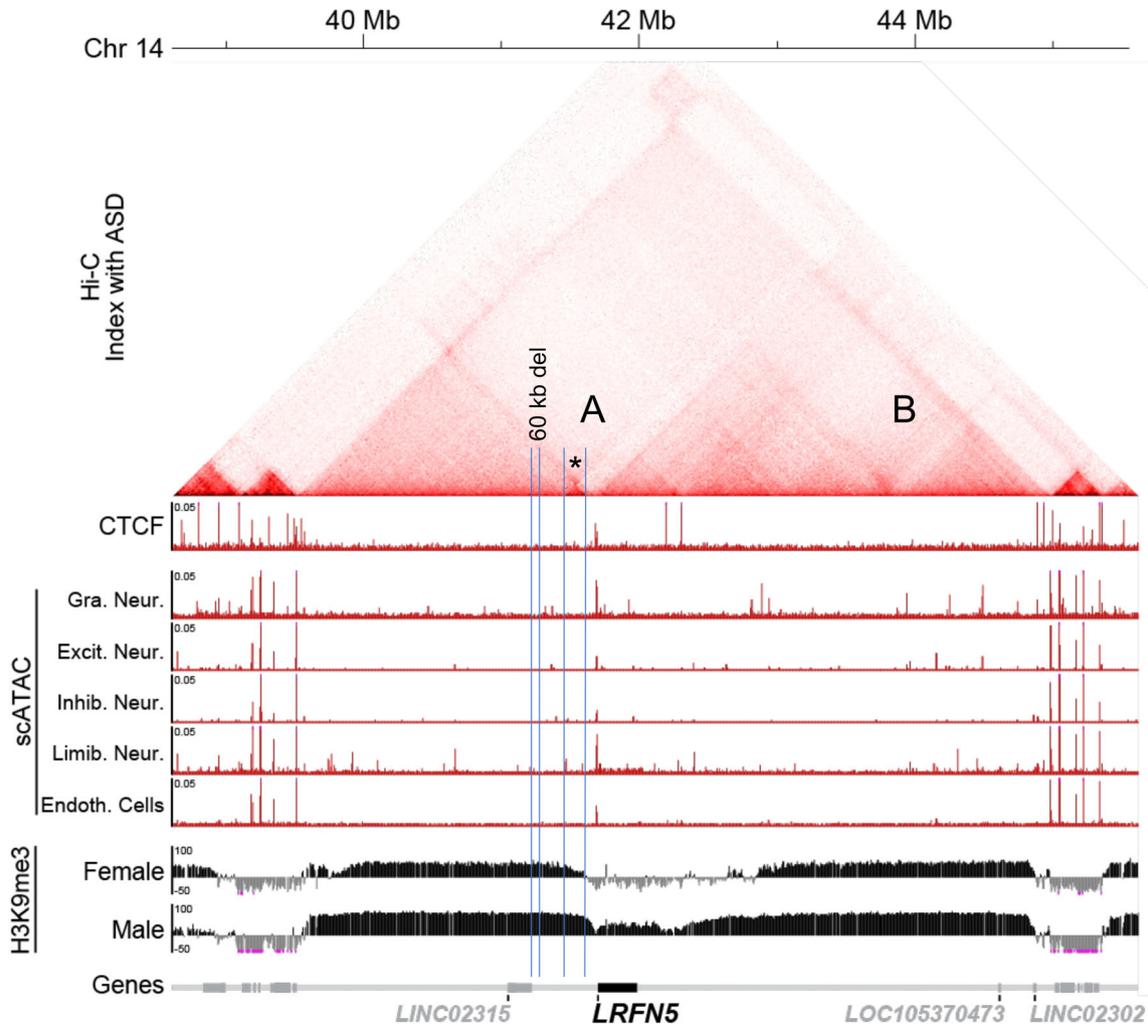


FIGURE 4 Top: The capture-HiC result of the index male with ASD, showing three subTADs inside the 5.4 Mb mega-TAD. The small diamond marked with an asterisk indicates the 172 kb deletion, and the position of the common 60 kb deletion polymorphism is also shown. Middle part shows CTCF sites (marking e.g., TAD boundaries and the peak of the mega-TAD) and single-cell ATAC sites in selected cell types (indicating areas of open chromatin). On the bottom is the H3K9me3 chromatin profiles of the two individuals with the most pronounced male and female pattern from the control profiles displayed in Figure 2

difference. This shows that the difference in fibroblast H3K9me3 profiles at the *LRFN5* gene is not reflected in differences in CpG methylation of blood leukocyte DNA in the same region.

Allelic interaction must be frequent in the *LRFN5* locus

If the *LRFN5* locus structure is critical for brain function, how is it maintained when copy number changes are quite common? A possible answer to this question was unexpectedly found upon examining the allelic distribution of a common 60 kb deletion in the middle of the *LRFN5*-upstream gene desert (chr14:41,609,383-41,669,664 [hg19]), see Table 2 (position of deletion is indicated in Figure 4). We examined three large and independent cohorts of individuals (two

from Norway and one from the Netherlands) who had been investigated with a high-density SNP-array because of a developmental disorder, usually including variable degrees of developmental delay (DD). In addition, one Dutch cohort ($n = 1416$) of mainly medical students served as a control. The deletion's minor allele frequency (MAF) was around 8.5% in the Dutch population, 15% in the Norwegian population, and 14% in the gnomAD database of population variation (gnomad.broadinstitute.org) (Karczewski et al., 2020). When comparing the observed allelic distribution in these cohorts with the expected distribution as per Hardy-Weinberg equilibrium, there were too few heterozygotes in all cohorts. We found that the loss of heterozygote wild type (wt)/deletion (del) was of equal magnitude in both directions, that is, to wt/wt and to del/del. In the three patient cohorts, this loss was 20%, 24%, and 26%, respectively, but in the student cohort, it was 80% (Table 2). This

surprising difference was not a technical artifact. 12–18 SNP-array oligonucleotides covered the 60 kb deletion, and > 5 in a row is usually sufficient for deletion detection. Furthermore, if this was due to missed wt/del calling, the high number of del/del homozygotes would still be incompatible with Hardy–Weinberg equilibrium. We are not aware of any other potential explanations for these observations than an early mitotic allelic conversion event with a frequency of one in five among heterozygotes in the DD group, and four in five among heterozygotes in the student group. This is far beyond expectation. Probably such allelic interaction also occurred in the homozygous wt/wt and del/del groups, but that would not be registered by our copy number test. Maybe allelic interaction is needed to establish monoallelic expression. In human fetal brain cell cultures from the striatum, we did indeed find preliminary evidence indicating monoallelic *LRFN5* expression, but since these data come from a single experiment, they need confirmation.

DISCUSSION

Despite extensive research, it has remained elusive why autism, and especially the higher-functioning variants, is more common in males than in females (Gillberg et al., 2006; Zhao et al., 2007). Here we show that the synaptic regulation and maintenance gene *LRFN5*, situated in the middle of a conserved gene desert capable of forming a 5.4 Mb mega-TAD, may be one answer to this question. The three main reasons for this are (1) remotely related pairs of brothers with a similar form of higher-functioning autism who shared the same maternally inherited *LRFN5* locus haplotype (Table 1), (2) the sex-influenced differences in locus chromatin structure in fibroblasts from normal males and females (Figures 2 and 3), and (3) the allelic interaction that must take place between the *LRFN5* loci, probably in early embryonic development, as evidenced by the striking “symmetrical” deviation from Hardy–Weinberg equilibrium of a 60 kb deletion polymorphism (Table 2). All this points to complex and sex-influenced regulation of a synapse-related protein.

There are eight putative SOX9 and SRY binding sites flanking the *LRFN5* locus (Figure 2 and Figure S4; TFBS_conc track in the UCSC browser). If one only considers sites that are highly conserved among all the 59 registered mammals in the conservation track, two SOX9 sites remain in the centromeric part of the locus, and one SOX9 and two SRY sites remain in the telomeric part the locus. Hypothetically, these sites could have importance for generation of the observed male/female chromatin difference (Figures 2 and 3). The mechanism behind the locus interaction is more difficult to explain. We are unaware of frequent allelic interaction and locus conversion in any other part of the genome, for example, the

frequency of inter-chromosomal gene conversion in gene families with more than two alleles has been estimated to be around 0.2% (Benovoy & Drouin, 2009). Gene conversion events are more common towards the 3'UTR end of protein-coding genes, and this is believed to be due to RNA transcription aiding the process (Benovoy & Drouin, 2009). Maybe *LINC02315*, an RNA gene upstream of *LRFN5* apparently ending in the common 60 kb deletion polymorphism, has a similar role (Figure 4).

The most interesting question is why frequent locus homogenization by allelic conversions occurs. We hypothesize that this process is needed to establish monoallelic expression, advantageous for fine-tuning *LRFN5* expression. Given the high conversion frequencies, allelic interaction is probably the rule, and allelic conversion a consequence. An exploratory *LRFN5* expression study suggested that monoallelic expression occurred in primary cultures of human fetal neurons from the striatum (data not shown), but we do not know if this is always the case or if it is stochastic. The next question is why fine-tuning of *LRFN5* expression is so critical that the gene, on top of having a 1.9 kb 5'UTR encoded by exons 1 and 2, needs to be framed by 2–3 Mb of conserved, presumed regulatory, gene deserts. Maybe the difference in locus conversion frequencies between a student cohort (4 in 5 heterozygotes were converted to homozygosity) and three DD-cohorts (1 in 4–5 heterozygotes were converted to homozygosity) provides a clue (Table 2). Hypothetically, if *LRFN5* plays a role for synaptic memory by strengthening or weakening synaptic connections, then complex and fine-tuned expression regulation would be expected, and allelic interaction could be a part of this regulation. This fits with *LRFN5*'s role as a protein involved in synapse strength and dynamics (Lin et al., 2018; Mah et al., 2010; Morimura et al., 2006), and it could have relevance for the photographic memory of details that some autistic individuals may have.

The concept of autism-related risk haplotypes of the *LRFN5* locus, variable locus structure influenced by the individual's sex, and early allelic homogenization, fits well with the complex pattern of ASD inheritance (Zhao et al., 2007). In the families with the ASD-susceptible “A-haplotype” described here, all ASD males were born to carrier mothers, and one nonpenetrant male was also recorded (Table 1). Of note, this male was hemimethylated at rs144497930, that is, at the position of the major TAD-peak around the two CTCF binding sites in Figure 4 (Table S1). Other family members and control individuals were fully methylated at this CpG. Maybe the epigenetic pattern predisposing to autism was not established during his embryogenesis. In a “locus-interaction-determines-ASD-susceptibility” model, this makes sense and would also explain why rare deletions recorded in DECIPHER may be pathogenic despite inheritance from a presumed normal male or female parent

(Figure S2). Such deletions could interfere with allelic interaction and epigenetic regulation. It should also be noted that while many different deletions may be seen in the *LRFN5* locus (Figure S4), duplications are rare, also in DECIPHER (Figure S3). Finally, the whole *LRFN5* locus or the *LRFN5* gene itself can be disrupted or deleted, apparently without (additional) phenotypic consequences, as indicated from the DECIPHER database, a gnomAD pLI of 0.56 (Karczewski et al., 2020), and our own diagnostic genomic copy-number records. This is as expected if monoallelic expression is the rule.

In conclusion, our work suggests that disturbances of *LRFN5* regulation can cause autism in males, and that *LRFN5* locus structure is influenced by the sex of the individual, that is, the presence of a Y-chromosome or not. Our most striking and statistically solid finding is the indirect evidence for frequent locus interactions *in trans* with allelic conversions, and these data also suggest that *LRFN5* locus regulation could be linked to synaptic memory or other essential brain functions.

ACKNOWLEDGMENTS

We are most grateful to members of the two families with autistic brothers for having contributed blood samples and skin biopsies and for their interest in contributing to research. This work could not have been done without the expert technical assistance of Hilde E Rusaas, Lene Hjertnes and Atle Brendehaug. We are also grateful to Sofia Douzougou Houge for critical feedback on the manuscript. This work was supported by HelseVest grant #911459.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Gunnar Houge  <https://orcid.org/0000-0002-6102-1513>

REFERENCES

- Autism Spectrum Disorders Working Group of The Psychiatric Genomics, C. (2017). Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism*, 8, 21. <https://doi.org/10.1186/s13229-017-0137-9>
- Benovoy, D., & Drouin, G. (2009). Ectopic gene conversions in the human genome. *Genomics*, 93(1), 27–32. <https://doi.org/10.1016/j.ygeno.2008.09.007>
- Cappuccio, G., Attanasio, S., Alagia, M., Mutarelli, M., Borzone, R., Karali, M., Genesio, R., Mormile, A., Nitsch, L., Imperati, F., Esposito, A., Banfi, S., Del Giudice, E., & Brunetti-Pierri, N. (2019). Microdeletion of pseudogene chr14.232.a affects *LRFN5* expression in cells of a patient with autism spectrum disorder. *European Journal of Human Genetics*, 27(9), 1475–1480. <https://doi.org/10.1038/s41431-019-0430-5>
- Casey, J. P., Magalhaes, T., Conroy, J. M., Regan, R., Shah, N., Anney, R., Shields, D. C., Abrahams, B. S., Almeida, J., Bacchelli, E., Bailey, A. J., Baird, G., Battaglia, A., Berney, T., Bolshakova, N., Bolton, P. F., Bourgeron, T., Brennan, S., Cali, P., Correia, C., ... Ennis, S. (2012). A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Human Genetics*, 131(4), 565–579. <https://doi.org/10.1007/s00439-011-1094-6>
- Choi, Y., Nam, J., Whitcomb, D. J., Song, Y. S., Kim, D., Jeon, S., Um, J. W., Lee, S. G., Woo, J., Kwon, S. K., Li, Y., Mah, W., Kim, H. M., Ko, J., Cho, K., & Kim, E. (2016). *SALM5* trans-synaptically interacts with LAR-RPTPs in a splicing-dependent manner to regulate synapse development. *Scientific Reports*, 6, 26676. <https://doi.org/10.1038/srep26676>
- de Bruijn, D. R., van Dijk, A. H., Pfundt, R., Hoischen, A., Merckx, G. F., Gradek, G. A., Lybæk, H., Stray-Pedersen, A., Brunner, H. G., & Houge, G. (2010). Severe progressive Autism associated with two de novo changes: A 2.6-Mb 2q31.1 deletion and a balanced t(14;21)(q21.1;p11.2) translocation with long-range epigenetic silencing of *LRFN5* expression. *Molecular Syndromology*, 1(1), 46–57. <https://doi.org/10.1159/000280290>
- Devlin, B., & Scherer, S. W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development*, 22(3), 229–237. <https://doi.org/10.1016/j.gde.2012.03.002>
- Gillberg, C., Cederlund, M., Lamber, K., & Zeijlon, L. (2006). Brief report: "the autism epidemic". The registered prevalence of autism in a Swedish urban area. *Journal of Autism and Developmental Disorders*, 36(3), 429–435. <https://doi.org/10.1007/s10803-006-0081-6>
- Goto-Ito, S., Yamagata, A., Sato, Y., Uemura, T., Shiroshima, T., Maeda, A., Imai, A., Mori, H., Yoshida, T., & Fukai, S. (2018). Structural basis of trans-synaptic interactions between PTPdelta and SALMs for inducing synapse formation. *Nature Communications*, 9(1), 269. <https://doi.org/10.1038/s41467-017-02417-z>
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., Churchhouse, C., ... Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, 51(3), 431–444. <https://doi.org/10.1038/s41588-019-0344-8>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., ... MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809), 434–443. <https://doi.org/10.1038/s41586-020-2308-7>
- Ko, J., & Kim, E. (2007). Leucine-rich repeat proteins of synapses. *Journal of Neuroscience Research*, 85(13), 2824–2832. <https://doi.org/10.1002/jnr.21306>
- Ko, J., Kim, S., Chung, H. S., Kim, K., Han, K., Kim, H., Jun, H., Kaang, B. K., & Kim, E. (2006). *SALM* synaptic cell adhesion-like molecules regulate the differentiation of excitatory synapses. *Neuron*, 50(2), 233–245. <https://doi.org/10.1016/j.neuron.2006.04.005>
- Kragesteen, B. K., Spielmann, M., Paliou, C., Heinrich, V., Schopflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A. M., Jerković, I., Harabula, I., Guckelberger, P., Pechstein, M., Wittler, L., Chan, W. L., Franke, M., Lupicñez, D. G., Kraft, K., Timmermann, B., Vingron, M., ... Andrey, G. (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nature Genetics*, 50(10), 1463–1473. <https://doi.org/10.1038/s41588-018-0221-x>
- Lauritsen, M. B., Pedersen, C. B., & Mortensen, P. B. (2005). Effects of familial risk factors and place of birth on the risk of autism: A nationwide register-based study. *Journal of Child Psychology and Psychiatry*, 46(9), 963–971. <https://doi.org/10.1111/j.1469-7610.2004.00391.x>
- Lin, Z., Liu, J., Ding, H., Xu, F., & Liu, H. (2018). Structural basis of *SALM5*-induced PTPdelta dimerization for synaptic differentiation. *Nature Communications*, 9(1), 268. <https://doi.org/10.1038/s41467-017-02414-2>

- Mah, W., Ko, J., Nam, J., Han, K., Chung, W. S., & Kim, E. (2010). Selected SALM (synaptic adhesion-like molecule) family proteins regulate synapse formation. *The Journal of Neuroscience*, *30*(16), 5559–5568. <https://doi.org/10.1523/JNEUROSCI.4839-09.2010>
- Morimura, N., Inoue, T., Katayama, K., & Aruga, J. (2006). Comparative analysis of structure, expression and PSD95-binding capacity of Lrfn, a novel family of neuronal transmembrane proteins. *Gene*, *380*(2), 72–83. <https://doi.org/10.1016/j.gene.2006.05.014>
- Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., Mandell, D. S., Miller, L. A., Pinto-Martin, J., Reaven, J., Reynolds, A. M., Rice, C. E., Schendel, D., & Windham, G. C. (2007). The epidemiology of autism spectrum disorders. *Annual Review of Public Health*, *28*, 235–258. <https://doi.org/10.1146/annurev.publhealth.28.021406.144007>
- Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T. R., Correia, C., Abrahams, B. S., Almeida, J., Bacchelli, E., Bader, G. D., Bailey, A. J., Baird, G., Battaglia, A., Berney, T., Bolshakova, N., Bölte, S., Bolton, P. F., ... Betancur, C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, *466*(7304), 368–372. <https://doi.org/10.1038/nature09146>
- Seabold, G. K., Wang, P. Y., Chang, K., Wang, C. Y., Wang, Y. X., Petralia, R. S., & Wenthold, R. J. (2008). The SALM family of adhesion-like molecules forms heteromeric and homomeric complexes. *The Journal of Biological Chemistry*, *283*(13), 8395–8405. <https://doi.org/10.1074/jbc.M709456200>
- Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y. H., Hicks, J., Spence, S. J., Lee, A. T., Puura, K., Lehtimäki, T., Ledbetter, D., ... Wigler, M. (2007). Strong association of de novo copy number mutations with autism. *Science*, *316*(5823), 445–449. <https://doi.org/10.1126/science.1138659>
- Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J. T., Abrahams, B. S., Salyakina, D., Imielinski, M., Bradfield, J. P., Sleiman, P. M., Kim, C. E., Hou, C., Frackelton, E., Chiavacci, R., Takahashi, N., Sakurai, T., Rappaport, E., Lajonchere, C. M., Munson, J., Estes, A., ... Hakonarson, H. (2009). Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*, *459*(7246), 528–533. <https://doi.org/10.1038/nature07999>
- Weiss, L. A., Shen, Y., Korn, J. M., Arking, D. E., Miller, D. T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M. A., Green, T., Platt, O. S., Ruderfer, D. M., Walsh, C. A., Altshuler, D., Chakravarti, A., Tanzi, R. E., Stefansson, K., Santangelo, S. L., Gusella, J. F., Sklar, P., ... Autism, C. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *The New England Journal of Medicine*, *358*(7), 667–675. <https://doi.org/10.1056/NEJMoa075974>
- Yu, M., & Ren, B. (2017). The three-dimensional Organization of Mammalian Genomes. *Annual Review of Cell and Developmental Biology*, *33*, 265–289. <https://doi.org/10.1146/annurev-cellbio-100616-060531>
- Zhao, X., Leotta, A., Kustanovich, V., Lajonchere, C., Geschwind, D. H., Law, K., Law, P., Qiu, S., Lord, C., Sebat, J., Ye, K., & Wigler, M. (2007). A unified genetic theory for sporadic and inherited autism. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(31), 12831–12836. <https://doi.org/10.1073/pnas.0705803104>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Lybæk, H., Robson, M., de Leeuw, N., Hehir-Kwa, J. Y., Jeffries, A., Haukanes, B. I., Berland, S., de Bruijn, D., Mundlos, S., Spielmann, M., & Houge, G. (2022). *LRFN5* locus structure is associated with autism and influenced by the sex of the individual and locus conversions. *Autism Research*, *15*(3), 421–433. <https://doi.org/10.1002/aur.2677>