The complex etiology of autism spectrum disorder due to missense mutations of CHD8

白石, 大智

ARTICLE

The complex etiology of autism spectrum disorder due to missense mutations of *CHD8*

Taichi Shiraishi[1], Yuta Katayama[1], Masaaki Nishiyama[2], Hirotaka Shoji[3], Tsuyoshi Miyakawa[3], Taisuke Mizoo[1], Akinobu Matsumoto[4], Atsushi Hijikata[5], Tsuyoshi Shirai[6], Kouta Mayanagi[7], and Keiichi I. Nakayama[1,8]*

[1]Division of Cell Biology, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Fukuoka, Fukuoka 812-8582, Japan
[2]Department of Histology and Cell Biology, Graduate School of Medical Sciences, Kanazawa University, Kanazawa 920-8640, Japan
[3]Division of Systems Medical Science, Center for Medical Science, Fujita Health University, Toyoake, Aichi 470-1192, Japan
[4]Division of Biological Science, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan
[5]School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan
[6]Department of Computer Bioscience, Nagahama Institute of Bio-Science and Technology, 1266 Tamura-Cho, Nagahama, Shiga 526-0829, Japan
[7]Department of Drug Discovery Structural Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Fukuoka, Fukuoka 812-8582, Japan
[8]Anticancer Strategies Laboratory, TMDU Advanced Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan

*email: nakayama.keiichi@tmd.ac.jp

CHD8 is an ATP-dependent chromatin-remodeling factor encoded by the most frequently mutated gene in individuals with autism spectrum disorder (ASD). Although many studies have examined the consequences of *CHD8* haploinsufficiency in cells and mice, few have focused on missense mutations, the most common type of *CHD8* alteration in ASD patients. We here characterized *CHD8* missense mutations in ASD patients according to six prediction scores and experimentally examined the effects of such mutations on the biochemical activities of CHD8, neural differentiation of embryonic stem cells, and mouse behavior. Only mutations with high prediction scores gave rise to ASD-like phenotypes in mice, suggesting that not all *CHD8* missense mutations detected in ASD patients are directly responsible for the development of ASD. Furthermore, we found that mutations with high scores cause ASD by mechanisms either dependent on or independent of loss of chromatin-remodeling function. Our results thus provide insight into the molecular underpinnings of ASD pathogenesis caused by missense mutations of *CHD8*.

INTRODUCTION

Autism spectrum disorder (ASD) is a prevalent neurodevelopmental disorder primarily defined by impaired social interaction and communication as well as by restricted and repetitive behaviors or interests [1]. Given that genetic factors have been strongly implicated in the etiology of ASD, several large-scale genomic screening projects have recently been performed for individuals with this condition [2-7]. The most frequently mutated gene was found to be *CHD8*, which encodes chromodomain helicase DNA-binding protein 8 [8-11] and is currently attracting much attention as one of the most promising candidate genes to play a role in the pathogenesis of ASD. Whereas homozygous deletion of *Chd8* in mice results in early embryonic death [12-14], mice heterozygous for *Chd8* deletion manifest autistic-like phenotypes such as altered social behavior, increased anxiety-like behavior, and cognitive deficits [14-21]. CHD8-haploinsufficient neural progenitor cells (NPCs) differentiated from mouse or human embryonic stem cells (ESCs) have also been found to show aberrant expression of neural differentiation–related genes [22-24]. Although many previous studies have indicated that CHD8 haploinsufficiency gives rise to ASD, many mutations that do not lead to haploinsufficiency, such as missense mutations, have also been identified in the *CHD8* gene. The molecular mechanisms by which such missense mutations might give rise to ASD have remained unknown.

Proteins of the CHD family are also members of the SNF2 superfamily of ATP-dependent chromatin-remodeling factors that regulate gene transcription by altering chromatin structure [25, 26]. The nine members of the CHD family that have been identified in mammals are classified into three subfamilies: Subfamily I includes CHD1 and CHD2; subfamily II comprises CHD3, CHD4, and CHD5; and subfamily III consists of CHD6, CHD7, CHD8, and CHD9 [27, 28]. In particular, members of subfamily III are characterized by the presence of two chromatin organization modifier (chromo) domains, a helicase/ATPase domain, a DNA binding domain, and a Brahma and Kismet (BRK) domain. The chromo domain is highly conserved among chromatin binding proteins [29, 30] and binds to methylated lysine residues of histones [31]. The helicase/ATPase domain is also highly conserved among chromatin-remodeling factors and supports the sliding movement of histones along DNA in a manner dependent on ATP hydrolysis [32]. The DNA binding domain of CHD8 contains a SANT-like but with several insertions (SLIDE) domain, which is thought to play a key role in binding to DNA [33, 34]. Lastly, the BRK domain is found only in chromatin-remodeling enzymes of metazoans and interacts with CCCTC-binding factor (CTCF) [35, 36].

CHD8 has been shown to possess ATP-dependent histone-sliding activity *in*

vitro [37, 38], suggesting that it regulates the accessibility of DNA to nuclear factors by altering chromatin structure *in vivo*. In addition, CHD8 has been found to accumulate at enhancer and promoter regions of genes by chromatin immunoprecipitation (ChIP) analysis [14, 15, 21, 23, 39]. These observations raise the possibility that aberrant chromatin-remodeling activity of CHD8 may result in widespread dysregulation of gene expression. Indeed, missense mutations of *CHD3* and *CHD7* that are responsible for a neurodevelopmental syndrome and CHARGE syndrome, respectively, were shown to disturb the chromatin-remodeling activity of the encoded proteins *in vitro* [40, 41], with such molecular dysfunction being thought to underlie the pathogenesis of these syndromes. However, whether *CHD8* missense mutations of ASD patients impair the chromatin-remodeling activity of CHD8 and whether such molecular dysfunction of CHD8 is directly responsible for the development of ASD have remained unknown.

We here examined the chromatin-remodeling function of CHD8 proteins with mutations identified in individuals with ASD, and we found that several such mutations indeed impaired the ATP-hydrolyzing and histone-sliding activities of CHD8. We further found that such aberrant molecular function of CHD8 was associated with ASD-like behavior in mice and disrupted the expression of neural differentiation–related genes in NPCs. Of note, a mutation that did not disrupt the chromatin-remodeling function of CHD8 but which induced ectopic binding of the protein to DNA in ESCs was also found to lead to the development of ASD-like behavior in mice. Our results thus provide comprehensive insight into the molecular basis of the etiology of ASD associated with missense mutations of *CHD8*.

MATERIALS & METHODS

Data analysis for CHD8 missense variants

The list of known missense variants of human CHD8 for score determination was retrieved from the SFARI database (https://gene.sfari.org) and Mutation@A Glance (https://mutation.nagahama-i-bio.ac.jp/at_a_glance/CHD8) [42], the latter of which integrates clinical relevance annotations based on ClinVar, OMIM and Orphanet [43-45] databases. Information on single nucleotide polymorphisms (SNPs) of human *CHD8* was obtained from ClinVar, dbSNP, GnomAD and ToMMo [43, 46-48] databases. Various scores for the missense and SNP variants that characterize their possible pathogenicity, evolutionary conservation, or molecular structural features were determined (Supplementary Table 1). CADD [49] and REVEL [50] scores were adopted as damage prediction metrics for estimation of pathogenic risk. mCADD provides an average of CADD scores at a given codon for all possible missense substitutions

including those observed. The Cons_CHD8 score was calculated from information entropy based on the frequency of amino acids at a given site of the multiple sequence alignment of CHD8 orthologs from vertebrates. The structural context of each amino acid residue was evaluated with reference to the predicted 3D model structure of human CHD8 generated by AlphaFold2 [51]. The pLDDT score (pLDDTAF2) indicates the confidence level of model prediction accuracy, which is generally higher for residues within structured domains and lower for those in disordered regions. The relative solvent accessibility (rASAAF2) of each residue in the model structure was computed by the method of Shrake and Rupley [52], and is generally lower for residues that engage in intramolecular interactions that stabilize protein structure.

Plasmids

Complementary DNA encoding human CHD8 was subcloned into the Gateway pENTR 2B Dual Selection Vector (Thermo Fisher Scientific). Point mutations of *CHD8* were introduced with the use of a QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies). Wild-type (WT) and mutant *CHD8* sequences were subcloned into the pDEST10 vector with the Gateway cloning system (Thermo Fisher Scientific).

Protein expression and purification

Recombinant baculoviruses containing WT or mutant *CHD8* sequences were produced according to the Bac-to-Bac Baculovirus Expression System manual (Thermo Fisher Scientific). For protein expression, Sf21 cells (2×10^7/100 ml) were cultured for 3 days at 27°C with shaking at 125 rpm before infection with recombinant baculoviruses encoding WT or mutant CHD8 for another 3 days. The cells were then harvested by centrifugation at 1750 × *g* for 10 min at 25°C, and the WT and mutant CHD8 proteins were purified as previously described [38], with minor modifications. The cell pellet was thus washed with ice-cold phosphate-buffered saline containing 200 μM phenylmethylsulfonyl fluoride (PMSF) and then again isolated by centrifugation at 17,000 × *g* for 1 min at 4°C. The resulting pellet was suspended in 5 ml of buffer A (20% glycerol, 20 mM Hepes-NaOH (pH 8.0), 500 mM NaCl, 1.5 mM $MgCl_2$, 200 μM EDTA, 200 μM PMSF, 1% Nonidet P-40, 10 mM 2-mercaptoethanol, 10 mM imidazole, aprotinin (20 μg/ml), leupeptin (10 μg/ml)) and incubated on ice for 10 min, after which the cell lysate was centrifuged at 17,000 × *g* for 30 min at 4°C and the resulting supernatant was mixed with ProBond Nickel-Chelating Resin (Thermo Fisher Scientific) that had been washed three times with 1 ml of buffer A. The mixture was incubated for 3 h at 4°C with rotation, after which the resin was collected and washed

twice with 10 ml of buffer B (same as buffer A but with a Nonidet P-40 concentration of 0.01% and imidazole concentration of 50 mM), once with 10 ml of buffer C (same as buffer B but with NaCl at 150 mM), and once with 5 ml of buffer D (same as buffer C but with imidazole at 100 mM). The resin was transferred to a Bond Elut C18 cartridge (Agilent Technologies) and subjected to elution with 1 ml of buffer E (same as buffer D but with imidazole at 400 mM). The eluate was divided into portions, flash-frozen in liquid nitrogen, and stored at –80°C until use. Protein concentration was determined with the Bradford assay (Bio-Rad). Samples (1 μg) of each recombinant protein were subjected to SDS–polyacrylamide gel electrophoresis (PAGE) on a 7.5% acrylamide gel (acrylamide:bis-acrylamide ratio of 29:1), and the gel was stained with Coomassie Brilliant Blue R (Sigma-Aldrich).

Reconstitution of nucleosomes for *in vitro* assays

DNA fragments for assays or nucleosome reconstitution were produced by the polymerase chain reaction (PCR) with a vector containing the Widom 601 sequence as a template and a fluorescently labeled oligonucleotide (IRDye800-GAGCTCGGAACACTATCCGACTG, Integrated DNA Technologies) and a nonlabeled oligonucleotide (CTGGAGAATCCCGGTGC) as forward and reverse primers, respectively. Amplicons were purified with NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). The DNA was assembled into mononucleosomes by standard salt dialysis with histone octamer (16-0001, EpiCypher).

EMSA

An electrophoretic mobility-shift assay (EMSA) was performed essentially as described previously [38], with minor modifications. WT or mutant CHD8 proteins (4, 8, 16, or 32 nM) were mixed with an equal volume of 4 nM IRDye-labeled DNA or nucleosomes in reaction buffer 1 (20 mM Hepes-NaOH (pH 8.0), 75 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 1 mM dithiothreitol (DTT), 0.1% Tween 20, bovine serum albumin (BSA, 0.1 mg/ml)) and incubated for 30 min at 25°C. The reaction mixtures were then directly subjected to 0.5× Tris-borate-EDTA (TBE) native PAGE on a 4% (29:1) gel at 100 V for 60 min (for DNA) or 90 min (for nucleosomes). The gel was then scanned with an Odyssey imaging system (LI-COR) for visualization of bands. The signals for the free and CHD8-bound DNA or nucleosome bands were measured with the use of Fiji.

NADH-coupled ATP hydrolysis assay

WT or mutant CHD8 proteins (40 nM) were incubated for 60 min at 30°C with 20 nM

IRDye-labeled DNA or nucleosomes in reaction buffer 2 (25 mM Hepes-KOH (pH 8.0), 50 mM KCl, 2 mM $MgCl_2$, 1 mM DTT, BSA (0.1 mg/ml)) supplemented with 1 mM ATP, 0.5 mM phosphoenolpyruvate, 0.025 mM NADH, pyruvate kinase (25 U/ml), and lactate dehydrogenase (25 U/ml). The concentration of NADH was monitored by measurement of absorbance at 340 nm with an EnSpire multiplate reader (Perkin Elmer).

REA assay

For assay of restriction enzyme accessibility (REA), WT or mutant CHD8 proteins (20 nM) were incubated for up to 40 min at 30°C with 2 nM IRDye-labeled nucleosomes in reaction buffer 3 (20 mM Hepes-NaOH (pH 8.0), 50 mM NaCl, 3 mM $MgCl_2$, 2 mM DTT, BSA (0.1 mg/ml)) supplemented with 2 mM ATP and MfeI-HF (2 U/ml, New England BioLabs). The reaction was terminated by the addition of an equal volume of reaction stop buffer (10 mM Hepes-NaOH (pH 8.0), 0.6% SDS, 40 mM EDTA, 5% glycerol, proteinase K (0.1 mg/ml, Roche)) and incubation for 20 min at 55°C, after which DNA was purified with the use of NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) and subjected to 0.5× TBE native PAGE on an 8% (29:1) gel at 100 V for 120 min. The gel was scanned with an Odyssey imaging system (LI-COR) for visualization of DNA bands. For quantitation, the signals corresponding to cut and uncut bands were measured with the use of Fiji.

Cell culture

Sf21 cells (Gibco) were cultured at 27°C in Sf-900 II SFM medium (Invitrogen) supplemented with 5% fetal bovine serum (FBS, Invitrogen), penicillin (100 U/ml), and streptomycin (100 U/ml). Mouse ESCs [53] were cultured as previously described on a feeder layer of mouse embryonic fibroblasts in KnockOut Dulbecco's modified Eagle's medium (Thermo Fisher Scientific) supplemented with KnockOut Serum Replacement (Thermo Fisher Scientific), 2 mM L-glutamine, 1× nonessential amino acids, penicillin (100 U/ml), streptomycin (100 U/ml), 0.05 mM 2-mercaptoethanol, and leukemia inhibitory factor (1000 U/ml, Sigma-Aldrich). Alternatively, mouse ESCs were cultured as previously described [54] without feeder cells in ES medium (Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate, 1× nonessential amino acids, penicillin (100 U/ml), streptomycin (100 U/ml), 0.1 mM 2-mercaptoethanol, leukemia inhibitory factor (1000 U/ml), 1 μM PD0325901 (Cayman Chemical), and 3 μM CHIR99021 (Axon Medchem BV)).

Generation of mutation knock-in (KI) ESCs

For establishment of mouse ESCs heterozygous or homozygous for CHD8 mutations, the cells were transfected with a bicistronic CRISPR vector encoding Cas9 and a targeting single guide RNA (pX330, Addgene), single-stranded DNA or a vector containing the mutated *Chd8* sequence (Supplementary Table 2), and the pPGK-Puro vector (Addgene) with the use of the FuGENE HD Transfection Reagent (Promega). The transfected cells were subjected to selection for 48 h in the presence of puromycin (0.5 μg/ml), expanded for 5 to 7 days, and picked up as single colonies. ESC clones containing heterozygous or homozygous insertions were identified by PCR and restriction enzyme reaction (Supplementary Table 2). For the establishment of *Chd8*$^{+/L1072F}$ ESCs, an ESC cell line with the mutant (L1072F) allele and a deletion allele at the *Chd8* locus was transfected again with pX330, a single-stranded DNA oligonucleotide, and pPGK-Puro to rescue the deletion allele (Supplementary Table 2).

Generation of mutation KI mice

ESCs heterozygous for each CHD8 mutation were microinjected into C57BL/6J mouse blastocysts, which were then implanted into pseudopregnant ICR female mice. The resulting male chimeras were mated with C57BL/6J females, and germ-line transmission of the mutant allele was confirmed by PCR and restriction enzyme reaction, as for generation of mutation KI ESCs (Supplementary Table 2).

Animals

All animal experiments were approved by the animal ethics committee of Kyushu University (A22-194-0) and were conducted in compliance with the university guidelines and regulations for animal experimentation. All mice were housed in the specific pathogen–free animal facility at Kyushu University in accordance with institutional guidelines under the following conditions: ambient temperature of 22°C, 50% to 60% humidity, 12-h-dark and 12-h-light cycle, and free access to water and rodent chow (CA-1, CLEA Japan).

General protocol for behavioral tests

Chd8 mutant or control mice were group-housed (four animals per cage) in a room with a 12-h-light and 12-h-dark cycle (lights on at 7:00 a.m.) and with free access to food and water. Behavioral tests were performed with mice at 9 to 17 weeks of age between 9:00 a.m. and 6:00 p.m. as described previously [14], unless indicated otherwise. Each apparatus was cleaned with dilute sodium hypochlorite solution before testing of each

animal in order to prevent bias due to olfactory cues.

Motor function tests

A wire hang test apparatus (O'Hara & Co.) was used to assess balance and grip strength. The apparatus consists of a box (21.5 by 22 by 23 cm) with a wire mesh grid (10 by 10 cm) on top that can be inverted. Mice at 9 to 12 weeks of age were placed individually on the wire mesh, which was then inverted, causing the animal to grip the wire. The latency to the mouse falling was recorded with a cutoff time of 60 s. Motor coordination and balance were tested with the rotarod test. Mice at 10 to 17 weeks of age were placed on a rotating drum with a diameter of 3 cm (Accelerating Rotarod, UGO Basile), and the time that each animal was able to maintain its balance on the rod during its acceleration from 4 to 40 rpm over 5 min was measured.

Open field test

Each mouse at 10 to 13 weeks of age was placed in the corner of an open field apparatus (40 by 40 by 30 cm, AccuScan Instruments) that was illuminated at 100 lux. A VersaMax system (AccuScan Instruments) with photobeam sensors was used for automatic storage of activity data. Total distance traveled, vertical activity (rearing, measured by counting the number of photobeam interruptions), and time spent in the central area (20 by 20 cm) were recorded over 120 min.

Light-dark transition test

The apparatus for the light-dark transition test consisted of a cage (21 by 42 by 25 cm) that was divided into two sections of equal size by a partition with a door (O'Hara & Co.). One chamber was made of white plastic and brightly illuminated (390 lux), whereas the other was black and dark (2 lux). Mice at 9 to 13 weeks of age were placed in the dark side and allowed to move freely between the two chambers with the door open for 10 min. The number of transitions between the two compartments, latency to the first entry into the light chamber, distance traveled, and time spent in each chamber were recorded with the use of ImageLD software.

Elevated plus-maze test

The apparatus consisted of two open arms (25 by 5 cm) and two enclosed arms of the same size with 15-cm-high transparent walls (O'Hara & Co.). The arms and central square were made of white plastic plates and were elevated to a height of 55 cm above the floor. The likelihood of animals falling from the apparatus was minimized by the

presence of 3-mm-high plastic ledges on the open arms. Arms of the same type were arranged on opposite sides. Each mouse at 10 to 15 weeks of age was placed in the central square of the maze (5 by 5 cm) facing one of the closed arms, and its behavior was recorded over 10 min. The number of entries into and the time spent in the open and enclosed arms as well as distance traveled were measured with the use of ImageEP software.

Social interaction test in a novel environment

In the social interaction test, two mice at 10 to 16 weeks of age and of the same genotype that had previously been housed in different cages were placed together in a box (40 by 40 by 30 cm) and allowed to explore freely for 10 min. Analysis was performed automatically with the use of ImageSI software. The total number of contacts, total contact duration, mean duration per contact, and total duration of active contacts were measured. Images were captured at a rate of three frames per second, and the distance traveled between two successive frames was determined for each mouse. If the two mice contacted each other and the distance traveled by either mouse was >10 cm, then the behavior was considered an active contact. Active contacts included sniffing and following behavior.

ESC differentiation

Differentiation of mouse ESCs was induced essentially as described [54]. In brief, ESCs were deprived of feeder cells for three or four passages, after which they (1.4×10^6 cells) were allowed to form embryoid bodies during culture in nonadherent bacterial dishes for 8 days in CA medium (DMEM supplemented with 10% FBS, 2 mM L-glutamine, 1× nonessential amino acids, penicillin (100 U/ml), streptomycin (100 U/ml), and 0.1 mM 2-mercaptoethanol), with all-*trans* retinoic acid (5 μM) being added from day 4 to 8. For long-term culture of embryoid bodies in suspension, they were maintained in N-2 medium (DMEM/F12 (Thermo Fisher Scientific) supplemented with 2 mM L-glutamine, 1% BSA, penicillin (100 U/ml), streptomycin (100 U/ml), and 1× N-2 Supplement (Thermo Fisher Scientific)) for an additional 2 days and in Complete medium (DMEM supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, penicillin (100 U/ml), streptomycin (100 U/ml), 1× B-27 Supplement (Thermo Fisher Scientific), L-alanine (2 μg/ml), L-proline (7.76 μg/ml), $ZnSO_4$ (0.194 μg/ml), and vitamin B12 (0.34 μg/ml)) for 4 days.

Immunoblot analysis

Total protein was extracted from cells or brain tissue with lysis buffer (0.5% Triton X-100, 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM PMSF, 0.4 mM EDTA, 0.4 mM Na_3VO_4, 10 mM NaF, 10 mM sodium pyrophosphate, aprotinin (10 μg/ml), leupeptin (10 μg/ml)). Protein concentration was determined with a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). The extracts were subjected to SDS-PAGE on 5–20% ExtraPAGE One Precast Gels (Nacalai Tesque), the separated proteins were transferred to a polyvinylidene difluoride membrane, the membrane was incubated consecutively with primary antibodies to CHD8 (generated in-house) or to HSP90 (BD Biosciences) and horseradish peroxidase–conjugated secondary antibodies (Promega), and signals were visualized with SuperSignal West Pico PLUS (Thermo Fisher Scientific) reagents and a ChemiDoc imaging system (Bio-Rad).

RNA-seq analysis

Total RNA was extracted from cell lysates with the use of a PureLink RNA Mini Kit (Thermo Fisher Scientific), and the quality of the purified RNA was assessed with a 2100 Bioanalyzer (Agilent Technologies). After mRNA selection with a NEBNext Poly(A) mRNA Magnetic Isolation Module (New England BioLabs), libraries were prepared with a NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England BioLabs). The cDNAs were sequenced with a NovaSeq 6000 system (Illumina). The quality of the raw sequencing data was checked with FastQC (version 0.11.9), and trimming of adaptor sequences was performed with Trimmomatic (version 0.39). The total abundance of each transcript was calculated with the use of a series of programs including HISAT2 (version 2.2.0), featureCounts (version 2.0.1), and DESeq2 (version 1.36.0). RNA–sequencing (seq) reads were mapped against the mouse (mm10) genome. Gene ontology (GO) analysis of differentially expressed genes (false discovery rate (FDR) $q < 0.05$) was performed with the PANTHER classification system.

ChIP-seq analysis

ESCs were fixed with 1% formaldehyde, isolated by centrifugation at 2260 × *g* for 2 min at 4°C, flash-frozen in liquid nitrogen, and stored at –80°C until analysis. The cells were thawed on ice, suspended in ChIP buffer (5 mM Hepes-KOH (pH 8.0), 200 mM KCl, 1 mM $CaCl_2$, 1.5 mM $MgCl_2$, 5% sucrose, 0.5% Nonidet P-40, 10 μM cOmplete EDTA-free Protease Inhibitor Cocktail (Roche)), incubated for 10 min on ice, subjected to ultrasonic treatment with the use of a Diagenode Bioruptor, and digested with micrococcal nuclease for 40 min at 30°C. After the addition of EDTA to a final concentration of 0.1 mM, each digested sample was centrifuged at 15,000 × *g* for 10

min at 4°C, and the resulting supernatant was incubated on a rotary shaker overnight at 4°C with magnetic bead–conjugated antibodies to CHD8 (generated in-house). The beads were washed twice with ChIP buffer and once with Tris-EDTA buffer, after which bound proteins were eluted from the beads and cross-links were reversed by incubation overnight at 65°C with 1% SDS in Tris-EDTA buffer. DNA was purified with the use of a MinElute PCR Purification Kit (Qiagen). Libraries were constructed with a ThruPLEX DNA-Seq Kit (Takara Bio), and paired-end sequencing was performed with a NovaSeq 6000 system (Illumina). The quality of the raw sequencing data was checked with FastQC (version 0.11.9), and trimming of adaptor sequences was performed with Trimmomatic (version 0.39). The paired-end reads were mapped to the mm10 genome with bowtie2 (version 2.4.5), and duplicated reads were removed with Picard. Markedly enriched regions of the genome were identified with the MACS2 peak caller (version 2.2.7.1) after normalization of data with the use of S3norm. Peak annotation was performed with ChIPseeker (version 1.32.1) and BEDTools (version 2.30.0).

Statistical analysis

Quantitative data were analyzed with the two-tailed Student's *t* test, Welch's *t* test, the chi-square test, or two-way analysis of variance (ANOVA). Comparisons of each prediction score between *CHD8* SNPs of healthy individuals and *CHD8* missense mutations of ASD patients were performed with the Mann-Whitney *U* test. All statistical analysis was performed with the use of Excel or R software. A p value of <0.05 was considered statistically significant. The distribution of variants according to each prediction score was analyzed by principal component (PC) analysis with R software.

Data availability

All sequence data have been deposited in GEO under the accession number GSE227579.

RESULTS

Prediction of the extent of mutation-induced damage to the molecular function of CHD8

A total of 291 mutations of *CHD8* detected in individuals with ASD has been deposited in The Simons Foundation Autism Research Initiative (SFARI) database. Haploinsufficiency resulting from nonsense (25.1%) or frameshift (21.6%) mutations in *CHD8* is likely responsible for ASD pathogenesis, given that such genetic alterations

recapitulate ASD-like behavior in mice [14-21] and give rise to abnormal expression of neural differentiation–related genes in NPCs [22-24]. However, whereas missense mutations (35.1%) that give rise to only a single amino acid change are the most frequent type of *CHD8* mutation in ASD patients (Fig. 1A), few studies have focused on such mutations [55], leaving the molecular basis of ASD caused by these mutations unclear.

We adopted six independent parameters (CADD, mCADD, REVEL, Cons_CHD8, pLDDTAF2, and rASAAF2) for prediction of the extent of mutation-induced damage to the molecular function of CHD8 (Supplementary Table 1). PC analysis was performed in six dimensions for missense alterations identified as SNPs of *CHD8* in healthy individuals and for missense mutations of *CHD8* identified in ASD patients (Fig. 1B and Supplementary Fig. 1A). Only PC1 had significant power ($p = 2.0 \times 10^{-15}$) to discriminate between missense SNPs in healthy individuals and missense mutations in ASD patients (Fig. 1C). Three parameters (REVEL, mCADD, and Cons_CHD8) made substantial contributions to PC1 (Fig. 1D), and such scores for missense mutations in ASD patients were significantly higher than those for missense SNPs in healthy individuals (Fig. 1E and Supplementary Fig. 1B). These findings suggested that missense mutations of *CHD8* in ASD patients, especially those with high scores, have a larger impact on the molecular function of CHD8 compared with SNPs, but whether such mutations actually perturb the molecular function of CHD8 and which mutations are directly responsible for ASD pathogenesis remained to be elucidated.

We therefore next evaluated the molecular activity of CHD8 proteins with missense mutations found in ASD patients. Given that such mutations are widely distributed across the *CHD8* gene, we focused primarily on missense mutations identified within the known functional domains of the CHD8 protein (Fig. 1F). Seven of these missense mutations—V744I in the chromo domain [5]; L834P, R910Q, L1072F, and R1242Q in the helicase/ATPase domain [2, 4]; G1710V in the DNA binding domain [2]; and R2333C in the BRK domain [56, 57]—were selected as representatives in order to provide a broad review of mutations with various scores (Fig. 1G and Supplementary Fig. 1C-E). R1242Q and L834P mutations had high REVEL, mCADD, and Cons_CHD8 scores, whereas V744I, R910Q, G1710V, and R2333C mutations had low such scores and L1072F had only a high Cons_CHD8 score (Fig. 1G). In addition, K842R, a mutation not found in ASD patients but which impairs the helicase/ATPase activity of CHD8 [37], was included as a control for *in vitro* assays, and S2173X, a nonsense mutation located immediately upstream of the BRK domain that has been found in ASD patients, was also included in the present study.

Several mutations found in ASD patients impair the chromatin-remodeling function of CHD8

To evaluate experimentally the impact of missense mutations of CHD8, we produced WT and the above nine mutant CHD8 proteins with the use of the Bac-to-Bac system (Supplementary Fig. 2A) and performed several *in vitro* assays. First, WT and mutant CHD8 proteins were tested for their ability to bind to DNA or nucleosomes with the use of an EMSA. The binding ability of the mutant CHD8 proteins did not differ significantly from, or showed only a small decrease relative to, that of the WT protein (Fig. 2A and Supplementary Fig. 2C, D), indicating that CHD8 is able to bind to nucleosomes almost normally even with such mutations. Second, assessment of ATP hydrolysis activity revealed that two mutations (L834P and R1242Q) that are located in the helicase/ATPase domain and which showed high prediction scores, as well as the control mutation K842R, markedly impaired such activity of CHD8 (Fig. 2B and Supplementary Fig. 2E). Lastly, examination of histone-sliding activity with an REA showed that such activity was almost completely lost in the K842R and R1242Q mutants and was impaired to a lesser extent in the L834P and S2173X mutants (Fig. 2C and Supplementary Fig. 2F). On the other hand, the V744I, R910Q, L1072F, G1710V, and R2333C mutants did not differ significantly from the WT protein in this assay, which, with the exception of L1072F, is consistent with all the lower predicted damage scores for these mutations. A summary of the molecular profiles of the various CHD8 mutants (Fig. 2D) provides the first direct evidence that the chromatin-remodeling activity of CHD8 is indeed impaired by missense mutations associated with ASD.

Although we focused on missense mutations within the known functional domains of the CHD8 protein, several mutations located outside of these domains have been detected in ASD patients and some of these mutations have high prediction scores. In addition, although fewer in number, some missense SNPs in healthy individuals score as highly as do missense mutations in ASD patients (Fig. 1B, E). To examine whether such mutations with high scores also disrupt the molecular function of CHD8, we attempted to generate nine of the corresponding mutant CHD8 proteins and to assess their chromatin-remodeling activity *in vitro* (Supplementary Fig. 2B, G). Three of the mutant proteins (C1094R, H1097Q, and L1100P) were not produced in insect cells, however, suggestive of protein instability (Supplementary Fig. 2B). These mutations might therefore result in haploinsufficiency of CHD8 in humans, which can lead to the development of ASD. We performed the REA assay for the remaining six mutant proteins, which revealed that the Y1494H mutation, with high scores similar to those of

the L834P and R1242Q mutations, resulted in a moderate but significant loss of histone-sliding activity of CHD8 (Supplementary Fig. 2H). In contrast, the H859R, I1118V, and R1979H mutations, which had lower prediction scores similar to those of the R910Q and G1710V mutations, had no impact on the molecular function of CHD8. These results indicated that mutations with high prediction scores can give rise to loss of CHD8 function, regardless of whether they are located outside of functional domains or are identified in healthy individuals. On the other hand, the F1325S and R1889C mutations had high prediction scores but did not affect the molecular function of CHD8 (Supplementary Fig. 2G, H), suggestive of a discrepancy between the predicted damage scores and the actual measured activity for mutations located outside of functional domains.

The chromatin-remodeling activity of CHD8 is required for normal embryonic development in mice

To investigate the biological consequences of CHD8 molecular dysfunction, we generated KI mice with the R910Q, R1242Q, G1710V, or S2173X mutations. Given that mice with homozygous deletion of *Chd8* show early embryonic mortality [12-14], we first examined whether mice homozygous for each mutation also manifest such a phenotype. Mice homozygous for the R910Q or G1710V mutations, which essentially did not impair the histone-sliding activity of CHD8, were born according to a Mendelian ratio and developed without any apparent abnormalities (Fig. 3A, B). In contrast, mice homozygous for the R1242Q or S2173X mutations—which impaired both ATP hydrolysis and histone-sliding activities or histone-sliding activity alone, respectively—were not detected at either postnatal day (P) 21 (Fig. 3A) or embryonic day (E) 13.5 (Fig. 3B), suggesting that these mutations result in early embryonic mortality. To exclude the possibility that the R1242Q or S2173X mutations affect the expression level of CHD8, we examined CHD8 abundance in the mutant mouse brain at E13.5. Immunoblot analysis revealed that the amount of CHD8 in mice heterozygous for the R1242Q mutation (or in those heterozygous or homozygous for the R910Q or G1710V mutations) was similar to that in WT mice (Fig. 3C and Supplementary Fig. 3A), indicating that the embryonic lethality of the R1242Q mutation is due to disruption of the molecular function of CHD8 rather than to loss of protein expression. On the other hand, mice heterozygous for the S2173X mutation showed a reduced expression level of CHD8 compared with WT mice (Fig. 3C), suggesting that this mutation might affect CHD8 protein stability and give rise to embryonic mortality as a result of a reduced protein abundance.

Similar to mice with heterozygous deletion of *Chd8*, mice heterozygous for the four mutations examined did not manifest any apparent changes in general health as reflected by body weight, muscle strength, or locomotor activity, with the exception of a small increase in locomotor activity for R1242Q heterozygotes (Fig. 3D, F, G). With regard to brain weight, however, only mice heterozygous for the R1242Q or S2173X mutations—which resulted in early embryonic mortality in the homozygous state—manifested macrocephaly, similar to that evident in mice with heterozygous deletion of *Chd8* (Fig. 3E). Together, these results suggested that the chromatin-remodeling function of CHD8 plays an essential role in normal mouse development.

Molecular dysfunction of CHD8 results in ASD-like behavior in mice

We next performed behavioral tests to examine whether mice heterozygous for CHD8 missense mutations manifest behavioral abnormalities. We previously showed that mice heterozygous for deletion of *Chd8* show ASD-like behavioral phenotypes [14], including increased anxiety and abnormal social interaction. More specifically, with regard to anxiety, such mice spend less time in unsettling places such as the central area in the open field test, the light box in the light-dark transition test, and the open arms in the elevated plus-maze test. The anxiety-like phenotypes associated with heterozygous *Chd8* deletion in these tests were recapitulated by $Chd8^{+/-}$ mice in the present study, but no such behavioral abnormalities were detected for mice heterozygous for the R910Q or G1710V mutations of CHD8 (Fig. 4A–C and Supplementary Fig. 4). On the other hand, mice heterozygous for the R1242Q mutation, which gives rise to marked molecular dysfunction of CHD8 but does not affect its expression, showed a significantly reduced time spent at the center in the open field test as well as a reduced time spent in the light box and reduced number of transitions in the light-dark transition test compared with WT mice, indicative of increased anxiety-like behavior in the mutant animals (Fig. 4A, B and Supplementary Fig. 4A, B). Although the R1242Q mutant heterozygotes did not manifest significant differences from WT mice in the elevated plus-maze test, they tended to show abnormalities similar to those of mice heterozygous for *Chd8* deletion (Fig. 4C and Supplementary Fig. 4C). In addition, mice heterozygous for the S2173X mutation, which results in a reduced expression level of CHD8, manifested increased anxiety-like behavior in these various tests (Fig. 4A–C and Supplementary Fig. 4), thus largely recapitulating behavioral abnormalities found in *Chd8* haploinsufficient mice.

With regard to the abnormal social interaction characteristic of mice heterozygous for deletion of *Chd8*, these mice show no significant differences in the number of social contacts or the total duration of active contacts such as sniffing and

following behavior, whereas the total duration of contacts and the mean duration per contact are greatly increased (Fig. 4D and Supplementary Fig. 5A, B). Disorders of social communication in autistic patients include both a reduced frequency of communication due to reduced interest or concern as well as, conversely, excessive communication due to poor grasp of psychological distance from others. Similar defects in social communication have been detected not only in *Chd8* haploinsufficient mice but in a subset of other ASD mouse models [58, 59], which might reflect recapitulation of the characteristic subtypes of social interaction observed in individuals with ASD, such as aloofness, passivity, and active-but-odd behavior [60]. These defects in social interaction were reproduced in mice heterozygous for R1242Q or S2173X mutations (with the exception that the increase in mean duration per contact was not significant for S2173X heterozygotes) but not in those heterozygous for R910Q or G1710V (Fig. 4D and Supplementary Fig. 5A, B).

The above behavioral tests were performed with male mice. However, given recent reports of sexual dimorphism in *Chd8* haploinsufficient mice [18, 61], we performed representative behavioral tests on female mice with point mutations of *Chd8*. Similar to male mice, female mice heterozygous for the R1242Q or S2173X mutations manifested an increase in total contact time during the social interaction test (Supplementary Fig. 6A). Moreover, with regard to anxiety-like behavior in the elevated plus-maze test, such mutant female mice were highly similar to corresponding male animals. Female mice heterozygous for the S2173X mutation thus spent less time in the open arms in this test compared with WT controls, and those heterozygous for R1242Q showed a tendency to spend less time in the open arms and to travel shorter distances (Supplementary Fig. 6B). None of these abnormalities was apparent for female mice heterozygous for the R910Q or G1710V mutations (Supplementary Fig. 6), suggesting an overall lack of sexual dimorphism with regard to such behavior for the four mutant strains examined. Collectively, these results suggested that molecular dysfunction of CHD8 is associated with ASD-like behavioral phenotypes in both male and female mice.

Expression of neural differentiation–related genes is downregulated in *Chd8*$^{+/-}$ and *Chd8*$^{+/R1242Q}$ NPCs

Chd8$^{+/-}$ NPCs differentiated from ESCs were previously found to manifest abnormal expression of neural differentiation–related genes [22-24]. To elucidate how missense mutations of *Chd8* give rise to ASD-like behavior in mice, we first established mouse ESC lines heterozygous for *Chd8* deletion or for the R910Q or R1242Q mutations (Fig.

5A and Supplementary Fig. 7A), induced their differentiation into NPCs, and subjected the NPCs to RNA-seq analysis at three time points (days 0, 8, and 14) after the onset of neural induction (Fig. 5B). None of the mutant ESCs showed marked differences in gene expression compared with WT cells at day 0 (Fig. 5C, D), consistent with previous findings that CHD8 is not required for ESC maintenance and that the absence of CHD8 does not affect gene expression pattern in these cells [23, 24]. In contrast, the expression of 245 genes was downregulated in *Chd8*$^{+/-}$ NPCs compared with WT cells at day 14 (Fig. 5E). This pattern was recapitulated in NPCs heterozygous for the R1242Q mutation, which impairs the molecular function of CHD8 and gives rise to ASD-like behavior in mice, but not in NPCs heterozygous for the R910Q mutation (Fig. 5E). Of note, there was substantial overlap in the genes that were downregulated in *Chd8*$^{+/-}$ and *Chd8*$^{+/R1242Q}$ NPCs, although most of the downregulated genes were specific to each type of mutant NPC (Fig. 5F).

To clarify the function of the downregulated genes, we performed GO analysis. In addition to GO terms related to tissue development, many neural differentiation–related terms such as neurogenesis and nervous system development were commonly and significantly enriched among downregulated genes in *Chd8*$^{+/-}$ and *Chd8*$^{+/R1242Q}$ NPCs (Fig. 5G). This pattern was also observed with the results for NPCs at day 8 (Supplementary Fig. 7C, E). We further confirmed that the expression levels of individual genes included in the GO term for nervous system development, such as *Tbx3* and *Aldh1a2*, were commonly downregulated in *Chd8*$^{+/-}$ and *Chd8*$^{+/R1242Q}$ NPCs (Fig. 5H). Of note, it was previously shown that depletion of TBX3 impairs neuroepithelial differentiation of ESCs, and that a reduction in the activity of ALDH1A2, which plays a key role in retinoic acid metabolism in neurons, gives rise to ASD-like phenotypes in mice [62, 63]. Our results thus suggested that the chromatin-remodeling activity of CHD8 is essential for appropriate expression of neural differentiation–related genes in NPCs, and that disruption of the neural differentiation of NPCs due to molecular dysfunction of CHD8 might be central to the pathogenesis of ASD-like behavior in mice and humans.

The L1072F mutation of CHD8 maintains chromatin-remodeling activity but confers ASD-like behavior

Given that we found that missense mutations with high prediction scores indeed impaired the molecular function of CHD8 and gave rise to ASD-like behavioral phenotypes in mice, we next focused on the L1072F mutation, which also had relatively high scores for three parameters (Fig. 1G) but essentially did not impair the chromatin-

remodeling activity of CHD8 (Fig. 2D). To determine whether this mutation contributes to the development of ASD, we introduced it into mice and performed a series of behavioral tests. We found that male mice heterozygous for the L1072F mutation did not manifest any abnormalities in general health as reflected by body weight and locomotor activity (Supplementary Fig. 3C, D), but they showed macrocephaly and ASD-like behavior such as increased anxiety in the light-dark transition test, elevated plus-maze test, and open field test (Fig. 6A, B and Supplementary Figs. 3C, D and 4) as well as impaired social interaction in the social interaction test (Fig. 6C and Supplementary Fig. 5A, B). These results indicated that the L1072F mutation gives rise to the development of ASD-like phenotypes by a mechanism independent of the loss of chromatin-remodeling function of CHD8.

To identify this mechanism, we examined gene expression profiles of ESCs and NPCs heterozygous for the L1072F mutation by RNA-seq analysis. The number of differentially expressed genes was substantially greater for *Chd8*$^{+/L1072F}$ NPCs than for *Chd8*$^{+/-}$ or *Chd8*$^{+/R1242Q}$ cells, with the changes not being limited to downregulation but also including upregulation (Fig. 6D and Supplementary Fig. 7D). GO analysis revealed that terms associated with neuronal differentiation were enriched to a greater extent among downregulated genes than among upregulated genes (Fig. 6E and Supplementary Fig. 7E), suggesting that the expression of neural differentiation–related genes was attenuated in the *Chd8*$^{+/L1072F}$ mutant NPCs. Furthermore, such abnormal gene expression was observed even in ESCs at day 0 (Fig. 6F), suggesting that the L1072F mutation, unlike *Chd8* deletion or the R1242Q mutation, may affect the expression of genes in ESCs.

To investigate the mechanism by which the L1072F mutation gives rise to aberrant gene expression in ESCs, we generated ESCs homozygous for this mutation and examined whether the mutation might affect the binding of CHD8 to DNA by ChIP-seq analysis with antibodies to CHD8. ESCs homozygous for the R910Q mutation, which essentially did not disrupt the chromatin-remodeling activity of CHD8 or alter the pattern of gene expression, were also subjected to ChIP-seq analysis as a control. Neither mutation appeared to affect the stability or abundance of CHD8 protein in ESCs (Fig. 6G and Supplementary Fig. 7B). Consistent with previous studies [14, 15], we found that the WT protein bound specifically to the transcription start site (TSS) of genes, and a similar pattern was apparent for both CHD8 mutants (Fig. 6H). We then extracted genes to which CHD8 bound in this region, and found that the WT protein was associated with 4777 genes, whereas CHD8(R910Q) and CHD8(L1072F) bound to 5242 and 6893 genes, respectively (Fig. 6I). Whereas the CHD8 mutants with R910Q or

L1072F substitutions showed a slight reduction in binding to nucleosomes compared with the WT protein (Fig. 2A), the ChIP-seq analysis thus revealed that the number of target genes was increased for these mutant proteins (Fig. 6I). It is therefore possible that the reduced binding ability of the mutant proteins gives rise to a decrease in binding site selectivity, resulting in increased nonspecific binding to sites other than the intrinsic target sites. We found that 3538 genes were common targets for WT and both mutant CHD8 proteins, and that 1689 genes were specific targets for CHD8(L1072F) (Fig. 6I). GO analysis of these 1689 genes revealed that terms such as stem cell population maintenance and Wnt signaling pathway [64, 65] were significantly enriched (Fig. 6I). We confirmed that the expression levels of individual genes included in these GO terms were slightly but significantly downregulated in ESCs heterozygous for the L1072F mutation by RNA-seq analysis (Fig. 6J). These results suggested that the L1072F mutation may alter the binding specificity of CHD8 for DNA in ESCs, resulting in ectopic chromatin remodeling and changes in gene expression pattern. The aberrant gene expression apparent in NPCs heterozygous for the L1072F mutation is thus preceded by abnormal gene expression in ESCs and may give rise to the development of ASD-like phenotypes in mice. The expression level of CHD8 in mouse brain was not affected by the L1072F mutation (Supplementary Fig. 3A), and mice homozygous for this mutation were born according to Mendelian ratios and appeared to develop normally (Supplementary Fig. 3B), unlike mice homozygous for the R1242Q mutation, which died in utero, suggesting that the functions of CHD8 in embryogenesis and in ASD pathogenesis might differ.

DISCUSSION

Although missense mutations are the most common variant type for the *CHD8* gene in individuals with ASD, their effects on CHD8 molecular function and behavior have remained unclear. We here performed comprehensive molecular, cellular, and behavioral analyses to elucidate the molecular basis of ASD caused by missense mutations of *CHD8* (Fig. 7A). Specifically, the R1242Q mutation was shown to markedly impair the chromatin-remodeling function of CHD8 and expression of neural differentiation–related genes in NPCs as well as to give rise to ASD-like behavioral phenotypes in mice. In addition, the L834P mutation, which was also shown to disturb the chromatin-remodeling activity of CHD8, was previously found to affect locomotion and neurodevelopment in *Caenorhabditis elegans* [55]. Given that these two mutations have high REVEL, mCADD, and Cons_CHD8 scores, other missense mutations of *CHD8* with such high scores (including W751L, Y1168N, I1192T, D1222V, and

R1889H) may cause ASD in a similar manner.

In contrast, the R910Q and G1710V mutations were found to have essentially no effect on the chromatin-remodeling activity of CHD8, the transcriptome in NPCs, or mouse behavior, even though they have been identified in ASD patients. Indeed, ASD patients with the G1710V mutation do not manifest facial or gastrointestinal abnormalities, which are hallmark symptoms of such patients with mutations of *CHD8* [2]. These findings suggest that individuals with the R910Q or G1710V mutations are likely to develop ASD as a result of mutations in a different gene, and that not all missense mutations of *CHD8* are responsible for ASD development. Given that the mCADD and Cons_CHD8 scores for the R910Q and G1710V mutations were relatively low, it is possible that mutations with similar low scores (including I778V, R2035Q, G2277A, and S2437N) also may not directly contribute to the etiology of ASD.

Of note, we identified a missense mutation that did not affect the chromatin-remodeling function of CHD8 but which conferred ectopic DNA binding that was associated with an abnormal gene expression pattern in ESCs and NPCs as well as with ASD-like behavior in mice. These findings suggest that some mutations of *CHD8* may give rise to ASD by a mechanism independent of the loss of chromatin-remodeling activity, as appears to be the case for the L1072F mutation. Although our present results show that the L1072F mutation alters the binding regions of DNA targeted by CHD8, mutations of CHD8 might also affect its interactions with other proteins such as histone H1, β-catenin, p53, CHD7, and elongating RNA polymerase II [13, 66-68]. Another possibility is that CHD8 mutant proteins might interfere or compete with the WT protein, given that we analyzed gene expression profiles in ESCs and NPCs heterozygous for mutations and the fact that most *CHD8* mutations found in ASD patients are heterozygous. Some missense mutations of *CHD8* might thus contribute to the development of ASD by disrupting the interactions of CHD8 with DNA or proteins, in addition to those that impair chromatin-remodeling function. Although the mCADD and REVEL scores for L1072F were not especially high, its Cons_CHD8 score was. Mutations with similar scores (such as R812W, M838I, R952G, and R1580W) might therefore also be a cause of ASD.

With regard to sexual dimorphism, although there were almost no differences in behavior between male and female mice harboring the R910Q, R1242Q, G1710V, or S2173X mutations, female mice heterozygous for the L1072F mutation showed distinct outcomes compared with male mice (Supplementary Fig. 6). More specifically, male mice heterozygous for the L1072F mutation manifested ASD-like behavioral phenotypes including increased anxiety-like behavior and abnormal social interaction

(Fig. 6A–C), whereas female mice did not show any abnormalities in the elevated plus-maze test or in the social interaction test (Supplementary Fig. 6). These findings suggest that a mutation that impairs the chromatin-remodeling function of CHD8, such as the R1242Q mutation, might cause ASD in both males and females, whereas a mutation that contributes to the development of ASD by a mechanism independent of the loss of chromatin remodeling might cause ASD only in males, thereby giving rise to sexual dimorphism. We hypothesize that such a difference in pathogenic mechanisms might be responsible in part for the high prevalence of ASD in males [1], a notion that warrants further investigation with the use of clinical data from various human cohort studies.

It was unexpected that mutations found in healthy individuals resulted in a reduced abundance or dysfunction of the CHD8 protein (Supplementary Fig. 2B, H). This finding might suggest the existence of "hidden ASD patients" among supposedly healthy individuals. Given that ASD is a highly prevalent disorder [1] and that it is difficult to diagnose correctly when patients have relatively mild symptoms, it is possible that some individuals have certain ASD-like phenotypes but are not diagnosed with ASD. In the present study, the data regarding SNPs of *CHD8* were obtained in part from studies that did not focus on ASD, which might have resulted in the inclusion of undiagnosed ASD patients in the control group. Some mutations with high prediction scores supposedly identified in healthy individuals might therefore actually have been derived from such potential ASD patients. Nevertheless, we found that mutations in ASD patients generally scored higher than those in healthy individuals (Fig. 1E), indicative of the enrichment of more detrimental alterations of *CHD8* in ASD patients.

Given that our study was largely based on a mouse model, we limited our additional analysis to mouse ESCs rather than human ESCs. Considering the facts that all of the mutations examined in the present study are located at amino acids conserved between human and mouse CHD8 proteins and that homozygous deletion of the CHD8 gene has been shown to result in impaired neural differentiation in both mouse and human ESCs [22-24], we speculate that the introduction of the mutations examined here into human ESCs or induced pluripotent stem cells would result in neuronal differentiation defects similar to those indicated by our results. However, given that previous studies also suggest that there are differences in the genes whose expression levels are altered by CHD8 depletion in mouse versus human ESCs [22-24], we cannot exclude the possibility that the effects of CHD8 mutations on cell differentiation may differ between the two species. The establishment of a model that more closely resembles the pathology of ASD in humans may shed light on this issue.

In summary, our results indicate that the etiology of ASD due to *CHD8*

765 missense mutations is complex, with such mutations giving rise to ASD phenotypes
766 through different molecular mechanisms (Fig. 7B). Whether a mutation gives rise to
767 ASD might be predictable to some extent on the basis of its REVEL, mCADD, and
768 Cons_CHD8 scores.
769

ACKNOWLEDGEMENTS

We thank K. Tsunematsu and the Research Promotion Unit of the Medical Institute of Bioregulation at Kyushu University for technical assistance as well as A. Ohta for help with preparation of the manuscript. We also thank U. Kiyoe (Chiba University) for technical guidance on chromatin-remodeling assays. Computations were performed in part on the NIG supercomputer at ROIS National Institute of Genetics. This research was supported in part by KAKENHI grants from the Japan Society for the Promotion of Science (JSPS) to K.I.N. (JP23H00378) and to H.S. (JP16H06276 and JP22H04922), as well as by a grant from MEXT Promotion of Distinctive Joint Research Center Program to H.S. (JPMXP0621467949) and that from the Japan Agency for Medical Research and Development (AMED) to K.I.N. (23wm0425002).

AUTHOR CONTRIBUTIONS

T. Shiraishi performed experiments and wrote the manuscript. YK, MN, and AM conceived and supervised experimental design. T. Mizoo, HS, and T. Miyakawa performed behavioral tests for *Chd8* mutant mice. AH, T. Shirai, and KM performed analysis of *CHD8* mutations. KIN coordinated the study and wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at...

Correspondence and requests for materials should be addressed to Keiichi I. Nakayama.

REFERENCES

1. Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. *Nat. Rev. Genet.* 2008; **9**(5)**:** 341-355.
2. Bernier R, Golzio C, Xiong B, Stessman HA, Coe BP, Penn O *et al.* Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* 2014; **158**(2)**:** 263-276.
3. O'Roak BJ, Stessman HA, Boyle EA, Witherspoon KT, Martin B, Lee C *et al.* Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* 2014; **5:** 5595.
4. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 2014; **515**(7526)**:** 209-215.
5. D'Gama AM, Pochareddy S, Li M, Jamuar SS, Reiff RE, Lam AN *et al.* Targeted DNA sequencing from autism spectrum disorder brains implicates multiple genetic mechanisms. *Neuron* 2015; **88**(5)**:** 910-917.
6. Zhou X, Feliciano P, Shu C, Wang T, Astrovskaya I, Hall JB *et al.* Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. *Nat. Genet.* 2022; **54**(9)**:** 1305-1319.
7. Wang T, Hoekzema K, Vecchio D, Wu H, Sulovari A, Coe BP *et al.* Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. *Nat. Commun.* 2020; **11**(1)**:** 4932.
8. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 2012; **338**(6114)**:** 1619-1622.
9. Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A *et al.* Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* 2012; **149**(3)**:** 525-537.
10. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A *et al.* Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 2012; **485**(7397)**:** 242-245.
11. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012; **485**(7397)**:** 246-250.
12. Nishiyama M, Nakayama K, Tsunematsu R, Tsukiyama T, Kikuchi A, Nakayama KI. Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. *Mol. Cell. Biol.* 2004; **24**(19)**:** 8386-8394.

13. Nishiyama M, Oshikawa K, Tsukada Y, Nakagawa T, Iemura S, Natsume T *et al.* CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat. Cell. Biol.* 2009; **11**(2)**:** 172-182.
14. Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T *et al.* CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature* 2016; **537**(7622)**:** 675-679.
15. Platt RJ, Zhou Y, Slaymaker IM, Shetty AS, Weisbach NR, Kim JA *et al.* Chd8 Mutation Leads to Autistic-like Behaviors and Impaired Striatal Circuits. *Cell Rep.* 2017; **19**(2)**:** 335-350.
16. Gompers AL, Su-Feher L, Ellegood J, Copping NA, Riyadh MA, Stradleigh TW *et al.* Germline Chd8 haploinsufficiency alters brain development in mouse. *Nat. Neurosci.* 2017; **20**(8)**:** 1062-1073.
17. Suetterlin P, Hurley S, Mohan C, Riegman KLH, Pagani M, Caruso A *et al.* Altered neocortical gene expression, brain overgrowth and functional over-connectivity in Chd8 haploinsufficient mice. *Cereb. Cortex* 2018; **28**(6)**:** 2192-2206.
18. Jung H, Park H, Choi Y, Kang H, Lee E, Kweon H *et al.* Sexually dimorphic behavior, neuronal activity, and gene expression in Chd8-mutant mice. *Nat. Neurosci.* 2018; **21**(9)**:** 1218-1228.
19. Kawamura A, Katayama Y, Nishiyama M, Shoji H, Tokuoka K, Ueta Y *et al.* Oligodendrocyte dysfunction due to Chd8 mutation gives rise to behavioral deficits in mice. *Hum. Mol. Genet.* 2020; **29**(8)**:** 1274-1291.
20. Jimenez JA, Ptacek TS, Tuttle AH, Schmid RS, Moy SS, Simon JM *et al.* Correction to: Chd8 haploinsufficiency impairs early brain development and protein homeostasis later in life. *Mol. Autism* 2021; **12**(1)**:** 33.
21. Hulbert SW, Wang X, Gbadegesin SO, Xu Q, Xu X, Jiang YH. A novel Chd8 mutant mouse displays altered ultrasonic vocalizations and enhanced motor coordination. *Autism Res* 2020; **13**(10)**:** 1685-1697.
22. Wang P, Lin M, Pedrosa E, Hrabovsky A, Zhang Z, Guo W *et al.* CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in neurodevelopment. *Mol. Autism* 2015; **6:** 55.
23. Sood S, Weber CM, Hodges HC, Krokhotin A, Shalizi A, Crabtree GR. CHD8 dosage regulates transcription in pluripotency and early murine neural differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 2020; **117**(36)**:** 22331-22340.
24. Ding S, Lan X, Meng Y, Yan C, Li M, Li X *et al.* CHD8 safeguards early neuroectoderm differentiation in human ESCs and protects from apoptosis during

neurogenesis. *Cell Death Dis.* 2021; **12**(11)**:** 981.
25. Tyagi M, Imam N, Verma K, Patel AK. Chromatin remodelers: We are the drivers!! *Nucleus (Calcutta)* 2016; **7**(4)**:** 388-404.
26. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* 2017; **18**(7)**:** 407-422.
27. Hall JA, Georgel PT. CHD proteins: a diverse family with strong ties. *Biochem. Cell Biol.* 2007; **85**(4)**:** 463-476.
28. Marfella CG, Imbalzano AN. The Chd family of chromatin remodelers. *Mutat. Res.* 2007; **618**(1-2)**:** 30-40.
29. Tajul-Arifin K, Teasdale R, Ravasi T, Hume DA, Mattick JS, Group RG *et al.* Identification and analysis of chromodomain-containing proteins encoded in the mouse transcriptome. *Genome Res.* 2003; **13**(6B)**:** 1416-1429.
30. Yap KL, Zhou MM. Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription. *Biochemistry* 2011; **50**(12)**:** 1966-1980.
31. Min J, Zhang Y, Xu RM. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* 2003; **17**(15)**:** 1823-1828.
32. Hargreaves DC, Crabtree GR. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res.* 2011; **21**(3)**:** 396-420.
33. Ryan DP, Sundaramoorthy R, Martin D, Singh V, Owen-Hughes T. The DNA-binding domain of the Chd1 chromatin-remodelling enzyme contains SANT and SLIDE domains. *EMBO J.* 2011; **30**(13)**:** 2596-2609.
34. Zhong Y, Moghaddas Sani H, Paudel BP, Low JKK, Silva APG, Mueller S *et al.* The role of auxiliary domains in modulating CHD4 activity suggests mechanistic commonality between enzyme families. *Nat. Commun.* 2022; **13**(1)**:** 7524.
35. Allen MD, Religa TL, Freund SM, Bycroft M. Solution structure of the BRK domains from CHD7. *J. Mol. Biol.* 2007; **371**(5)**:** 1135-1140.
36. Ishihara K, Oshimura M, Nakao M. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol. Cell* 2006; **23**(5)**:** 733-742.
37. Thompson BA, Tremblay V, Lin G, Bochar DA. CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. *Mol. Cell. Biol.* 2008; **28**(12)**:** 3894-3904.
38. Manning BJ, Yusufzai T. The ATP-dependent chromatin remodeling enzymes CHD6, CHD7, and CHD8 exhibit distinct nucleosome binding and remodeling

activities. *J. Biol. Chem.* 2017; **292**(28)**:** 11927-11936.
39. Kawamura A, Katayama Y, Kakegawa W, Ino D, Nishiyama M, Yuzaki M *et al.* The autism-associated protein CHD8 is required for cerebellar development and motor function. *Cell Rep.* 2021; **35**(1)**:** 108932.
40. Snijders Blok L, Rousseau J, Twist J, Ehresmann S, Takaku M, Venselaar H *et al.* Author Correction: CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. *Nat. Commun.* 2019; **10**(1)**:** 2079.
41. Bouazoune K, Kingston RE. Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. *Proc. Natl. Acad. Sci. U. S. A.* 2012; **109**(47)**:** 19238-19243.
42. Hijikata A, Raju R, Keerthikumar S, Ramabadran S, Balakrishnan L, Ramadoss SK *et al.* Mutation@A Glance: an integrative web application for analysing mutations from human genetic diseases. *DNA Res.* 2010; **17**(3)**:** 197-208.
43. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014; **42**(Database issue)**:** D980-985.
44. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* 2015; **43**(Database issue)**:** D789-798.
45. Arroyo JI, Apperson SA, Cropp CB, Marafino BJ, Jr., Monath TP, Tesh RB *et al.* Effect of human gamma interferon on yellow fever virus infection. *Am. J. Trop. Med. Hyg.* 1988; **38**(3)**:** 647-650.
46. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001; **29**(1)**:** 308-311.
47. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020; **581**(7809)**:** 434-443.
48. Tadaka S, Hishinuma E, Komaki S, Motoike IN, Kawashima J, Saigusa D *et al.* jMorp updates in 2020: large enhancement of multi-omics data resources on the general Japanese population. *Nucleic Acids Res.* 2021; **49**(D1)**:** D536-D544.
49. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 2014; **46**(3)**:** 310-315.
50. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S *et al.*

REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* 2016; **99**(4)**:** 877-885.

51. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021; **596**(7873)**:** 583-589.
52. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* 1973; **79**(2)**:** 351-371.
53. Mishina M, Sakimura K. Conditional gene targeting on the pure C57BL/6 genetic background. *Neurosci. Res.* 2007; **58**(2)**:** 105-112.
54. Bibel M, Richter J, Lacroix E, Barde YA. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat. Protoc.* 2007; **2**(5)**:** 1034-1043.
55. Wong WR, Brugman KI, Maher S, Oh JY, Howe K, Kato M *et al.* Autism-associated missense genetic variants impact locomotion and neurodevelopment in Caenorhabditis elegans. *Hum. Mol. Genet.* 2019; **28**(13)**:** 2271-2281.
56. Kimura H, Wang C, Ishizuka K, Xing J, Takasaki Y, Kushima I *et al.* Identification of a rare variant in CHD8 that contributes to schizophrenia and autism spectrum disorder susceptibility. *Schizophr Res.* 2016; **178**(1-3)**:** 104-106.
57. McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y *et al.* De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol. Psychiatry* 2014; **19**(6)**:** 652-658.
58. Chao HT, Chen H, Samaco RC, Xue M, Chahrour M, Yoo J *et al.* Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* 2010; **468**(7321)**:** 263-269.
59. Spencer CM, Alekseyenko O, Serysheva E, Yuva-Paylor LA, Paylor R. Altered anxiety-related and social behaviors in the Fmr1 knockout mouse model of fragile X syndrome. *Genes Brain Behav.* 2005; **4**(7)**:** 420-430.
60. Scheeren AM, Koot HM, Begeer S. Social Interaction Style of Children and Adolescents with High-Functioning Autism Spectrum Disorder. *J Autism Dev Disord* 2012; **42**(10)**:** 2046-2055.
61. Cherepanov SM, Gerasimenko M, Yuhi T, Furuhara K, Tsuji C, Yokoyama S *et al.* Oxytocin ameliorates impaired social behavior in a Chd8 haploinsufficiency mouse model of autism. *BMC Neurosci.* 2021; **22**(1)**:** 32.
62. Esmailpour T, Huang T. TBX3 promotes human embryonic stem cell proliferation and neuroepithelial differentiation in a differentiation stage-dependent manner.

Stem Cells 2012; **30**(10)**:** 2152-2163.

63. Xu X, Li C, Gao X, Xia K, Guo H, Li Y *et al.* Excessive UBE3A dosage impairs retinoic acid signaling and synaptic plasticity in autism spectrum disorders. *Cell Res.* 2018; **28**(1)**:** 48-68.

64. Reya T, Clevers H. Wnt signalling in stem cells and cancer. *Nature* 2005; **434**(7035)**:** 843-850.

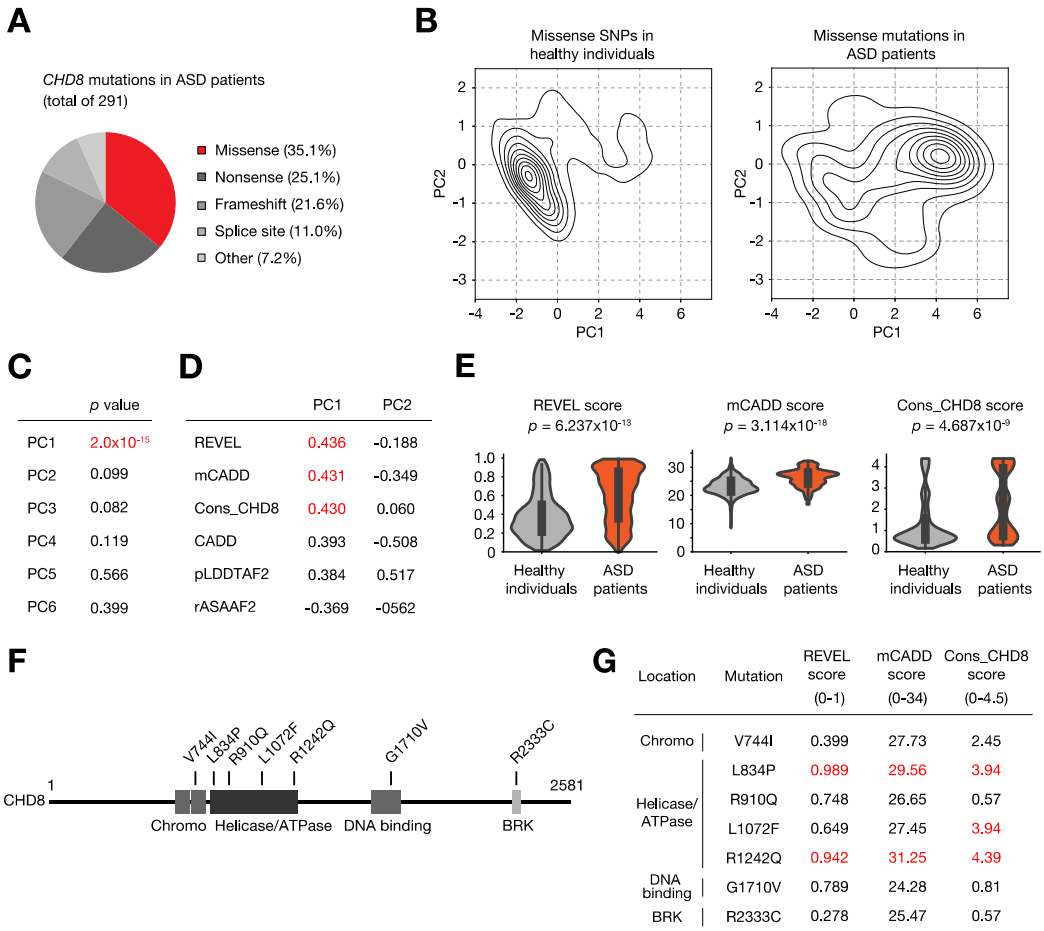
65. Nusse R. Wnt signaling and stem cell control. *Cell Res.* 2008; **18**(5)**:** 523-527.

66. Nishiyama M, Skoultchi AI, Nakayama KI. Histone H1 recruitment by CHD8 is essential for suppression of the Wnt-beta-catenin signaling pathway. *Mol. Cell. Biol.* 2012; **32**(2)**:** 501-512.

67. Batsukh T, Pieper L, Koszucka AM, von Velsen N, Hoyer-Fender S, Elbracht M *et al.* CHD8 interacts with CHD7, a protein which is mutated in CHARGE syndrome. *Hum. Mol. Genet.* 2010; **19**(14)**:** 2858-2866.

68. Rodriguez-Paredes M, Ceballos-Chavez M, Esteller M, Garcia-Dominguez M, Reyes JC. The chromatin remodeling factor CHD8 interacts with elongating RNA polymerase II and controls expression of the cyclin E2 gene. *Nucleic Acids Res.* 2009; **37**(8)**:** 2449-2460.

998 **FIGURE**

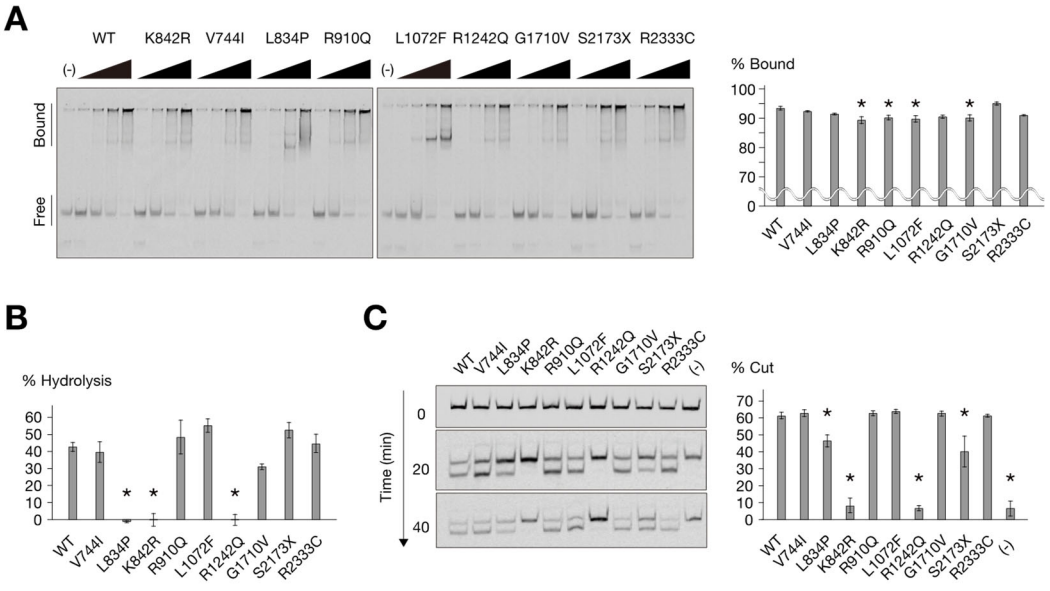


Shiraishi *et al.* Figure 1

999

1000 **Fig. 1 Missense mutations of *CHD8* in ASD patients show a score distribution**
1001 **distinct from that for SNPs of *CHD8* in healthy individuals. A** Categorization of
1002 ASD-associated *CHD8* mutations in the SFARI database. “Other” mutations include
1003 synonymous variants (3.1%), copy number variants (2.4%), in-frame variants (0.7%),
1004 translocation variants (0.7%), and intronic variants (0.3%). **B** Probability density
1005 distribution for *CHD8* variants shown on the PC1-PC2 plane. Missense SNPs in healthy
1006 individuals (left) and missense mutations in ASD patients (right) are plotted. **C** Biased
1007 distributions of SNPs and missense mutations among the PC axes determined by
1008 Wilcoxon tests. **D** Amplitude of factor loadings to PC1 and PC2. **E** Violin plots of
1009 parameter (REVEL, mCADD, and Cons_CHD8) scores for missense SNPs in healthy
1010 individuals and missense mutations in ASD patients. The *P* values were determined
1011 with the Mann-Whitney *U* test. **F** Schematic diagram of the human CHD8 protein

showing the domain organization and location of mutations examined in the present study. The 2581–amino acid protein contains four functional domains (chromo, helicase/ATPase, DNA binding, and BRK domains). **G** REVEL, mCADD, and Cons_CHD8 scores for *CHD8* missense mutations examined in the present study. Scores shown in red indicate that the mutation ranks in the top 25% of all *CHD8* missense mutations in ASD patients.

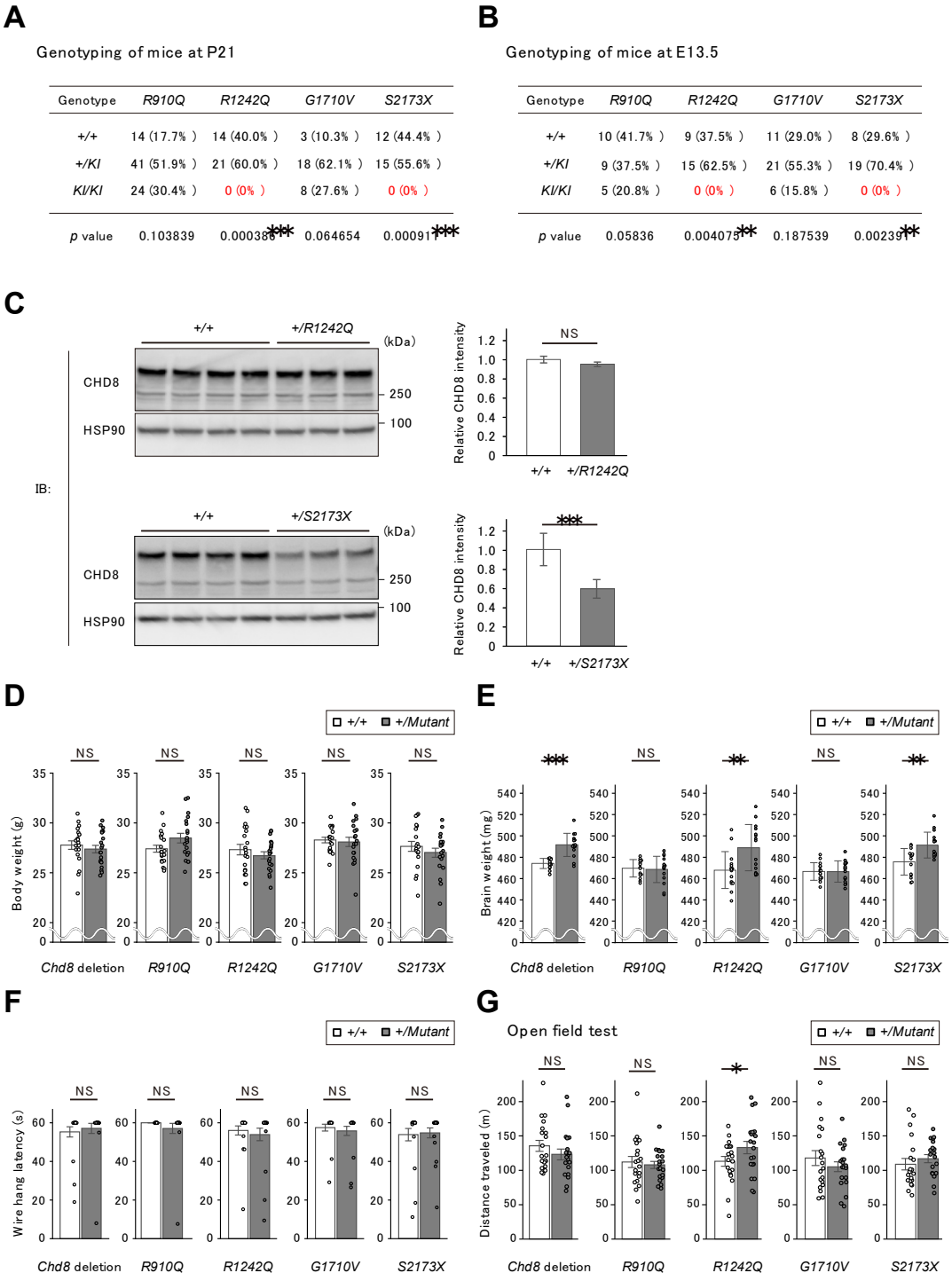


Mutation	V744I	L834P	R910Q	L1072F	R1242Q	G1710V	S2173X	R2333C
Functional domain	Chromo	Helicase	Helicase	Helicase	Helicase	DNA-bind		BRK
REVEL	0.399	0.989	0.748	0.649	0.942	0.789	-	0.278
mCADD	27.73	29.56	26.65	27.45	31.25	24.28	-	25.47
Cons_CHD8	2.45	3.94	0.57	3.94	4.39	0.81	-	0.57
Binding to nucleosomes	+	+	+	+	+	+	+	+
ATP hydrolysis	+	-	+	+	-	+	+	+
Histone sliding	+	±	+	+	-	+	±	+

Shiraishi *et al.* Figure 2

Fig. 2 The chromatin-remodeling activity of CHD8 is impaired by missense mutations associated with ASD. A Binding activity of WT and mutant CHD8 proteins measured by EMSA. IRDye-labeled nucleosomes (4 nM) were incubated with an equal volume of 4, 8, 16, or 32 nM CHD8 proteins for 30 min at 25ºC, after which the binding

mixtures were fractionated by 4% native PAGE. Representative gels (left panels) and quantification of bound nucleosome bands for samples with 32 nM WT or mutant CHD8 (right panel) are shown. **B** ATP hydrolysis activity of WT and mutant CHD8 proteins measured with an NADH-coupled assay in the presence of nucleosomes for 60 min at 30ºC. **C** Histone-sliding activity of WT and mutant CHD8 proteins measured with an REA assay. Cut and uncut DNA fragments were separated by 8% native PAGE at each time point during incubation at 30ºC. Representative gels (left) and quantification of cut DNA bands in samples at 40 min (right) are shown. All quantitative data in bar graphs are means ± SEM for three independent experiments. *$p < 0.05$ versus WT (Welch's *t* test). **D** Summary of the measured activities and prediction scores for WT and mutant CHD8 proteins, where + indicates activity equivalent to that of the WT protein, and ± and − indicate significantly reduced activity compared to that of the WT protein, 50% and 90% or 0% and 50%, respectively.

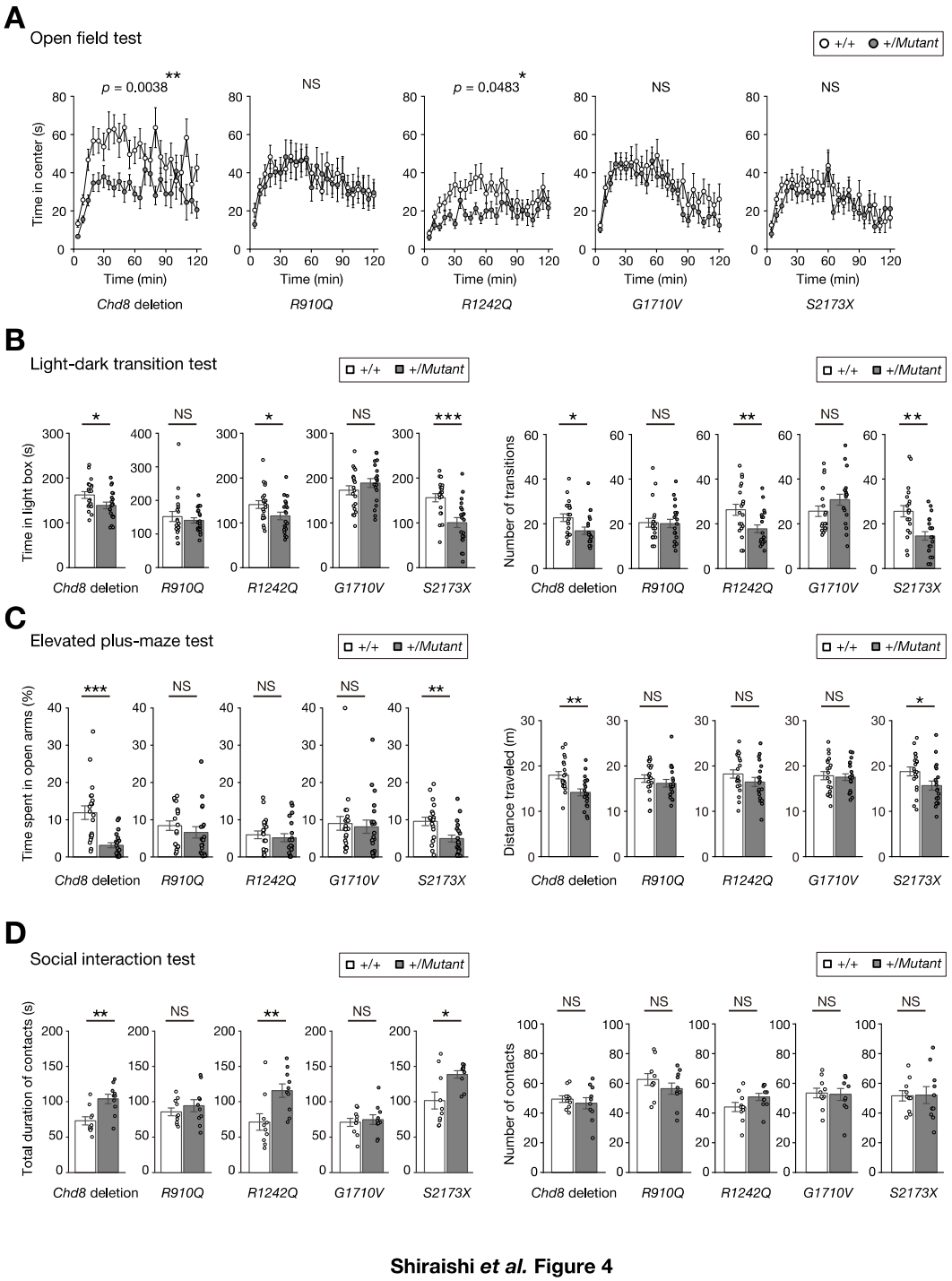


Shiraishi *et al.* Figure 3

Fig. 3 Molecular dysfunction of CHD8 gives rise to embryonic mortality in mice.

A, **B** Number and percentage of mice of the indicated genotypes at weaning (**A**) or at

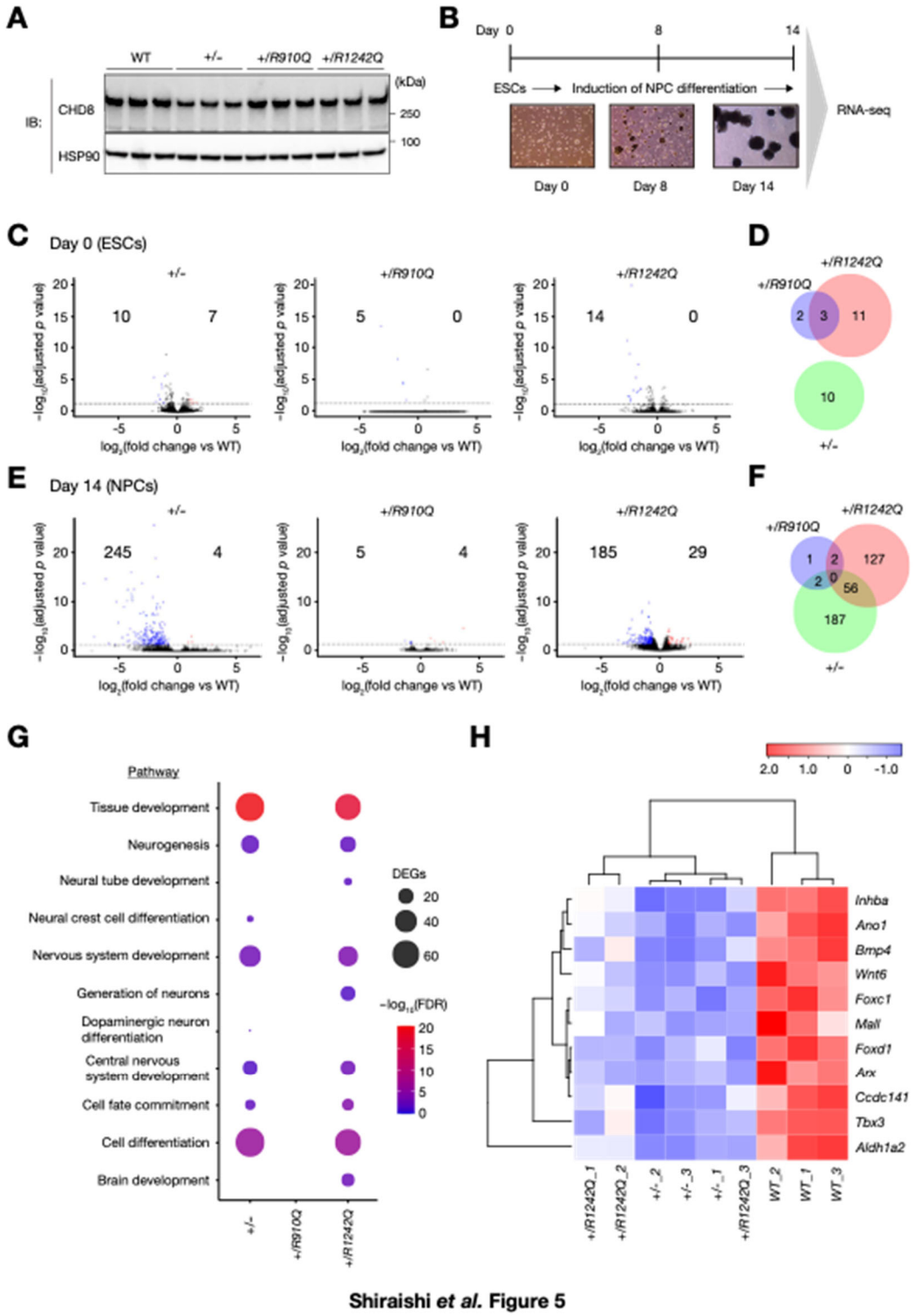
E13.5 (**B**) for offspring produced by mating male and female heterozygotes for each mutation. Genotypes are denoted by +/+ for WT and by +/*KI* and *KI*/*KI* for heterozygotes or homozygotes, respectively, for each mutation. *p* values were determined with the chi-square test. $^{**}p < 0.01$, $^{***}p < 0.001$. **C** Immunoblot (IB) analysis of CHD8 in the brain of mice of the indicated genotypes at E13.5. Representative data (left) and quantification of relative CHD8 intensity (right) are shown. HSP90 was examined as a loading control. $^{***}p < 0.001$; NS, not significant (unpaired two-tailed Student's *t* test). **D**–**G** Body weight (**D**), brain weight (**E**), wire hang latency in the wire hang test (**F**), and total distance traveled in the open field test (**G**) for adult male mice of the indicated genotypes. White bars indicate WT mice, and gray bars indicate littermates heterozygous for each mutation or deletion. $n = 20$ mice (**D**, **F**, **G**) or $n = 16$ mice (**E**) for each genotype. Quantitative data in bar graphs are means ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$; NS, not significant (unpaired two-tailed Student's *t* test).



Shiraishi *et al.* Figure 4

Fig. 4 Molecular dysfunction of CHD8 results in ASD-like behavior in mice. Time spent in the central area at each time point in the open field test (**A**), total time spent in the light chamber (left) and number of transitions between the light and dark chambers (right) in the light-dark transition test (**B**), percentage of time spent in the open arms

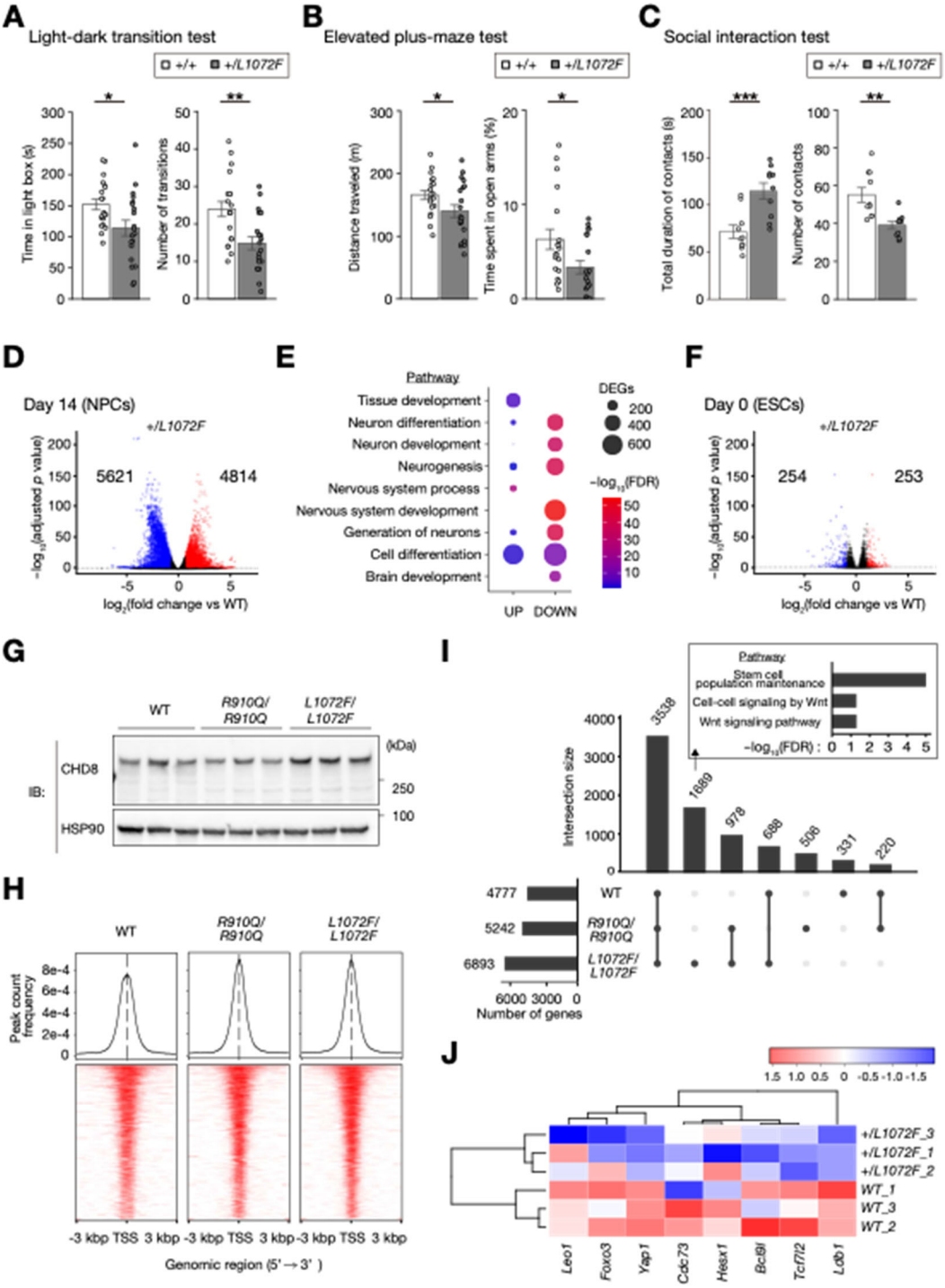
(left) and total distance traveled (right) in the elevated plus-maze test (**C**), and total duration of contacts (left) and total number of contacts (right) in the social interaction test (**D**) are shown for adult male mice of the indicated genotypes. White circles or bars indicate WT mice, and gray circles or bars indicate littermates heterozygous for each mutation or deletion ($n = 20$ for each genotype, with the exception that $n = 19$ for *R910Q* (+/+ and +/*R910Q*) and +/*R1242Q* in the elevated plus-maze test because one mouse fell from the apparatus in each case). All data are means ± SEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; NS, not significant (two-way ANOVA for repeated measures in **A**, and unpaired two-tailed Student's *t* test in **B** to **D**).



Shiraishi *et al.* Figure 5

Fig. 5 Neural differentiation–related gene expression is downregulated in both *Chd8*$^{+/-}$ and *Chd8*$^{+/R1242Q}$ NPCs. A Immunoblot analysis of CHD8 in mouse ESCs of

the indicated *Chd8* genotypes. **B** Schematic diagram of the procedure for inducing the differentiation of ESCs into NPCs. RNA-seq analysis was performed on samples collected at days 0, 8, and 14. The images are phase-contrast micrographs of cells at the corresponding times. **C** Volcano plots of RNA-seq data for *Chd8* mutant ESCs compared with WT ESCs at day 0. Significantly upregulated (FDR < 0.05, $\log_2$(fold change) > 1) or downregulated (FDR < 0.05, $\log_2$(fold change) < –1) genes in mutant versus WT cells are highlighted in red and blue, respectively. The numbers within the plots represent the numbers of differentially expressed genes (DEGs). **D** Venn diagram for the overlap of downregulated genes in mutant ESCs at day 0. **E** Volcano plots of RNA-seq data for *Chd8* mutant NPCs compared with WT NPCs at day 14. Significantly upregulated (FDR < 0.05, $\log_2$(fold change) > 0.7) or downregulated (FDR < 0.05, $\log_2$(fold change) < –0.7) genes in mutant versus WT cells are highlighted in red and blue, respectively. The numbers within the plots represent the numbers of differentially expressed genes. **F** Venn diagram for the overlap of downregulated genes in mutant NPCs at day 14. **G** GO analysis for downregulated genes in *Chd8* mutant NPCs at day14. **H** Heat map for the expression of representative genes included in the GO term of nervous system development and commonly downregulated in $Chd8^{+/-}$ and $Chd8^{+/R1242Q}$ NPCs at day 14. RNA-seq data for three biological replicates for each genotype are shown.



Shiraishi *et al.* Figure 6

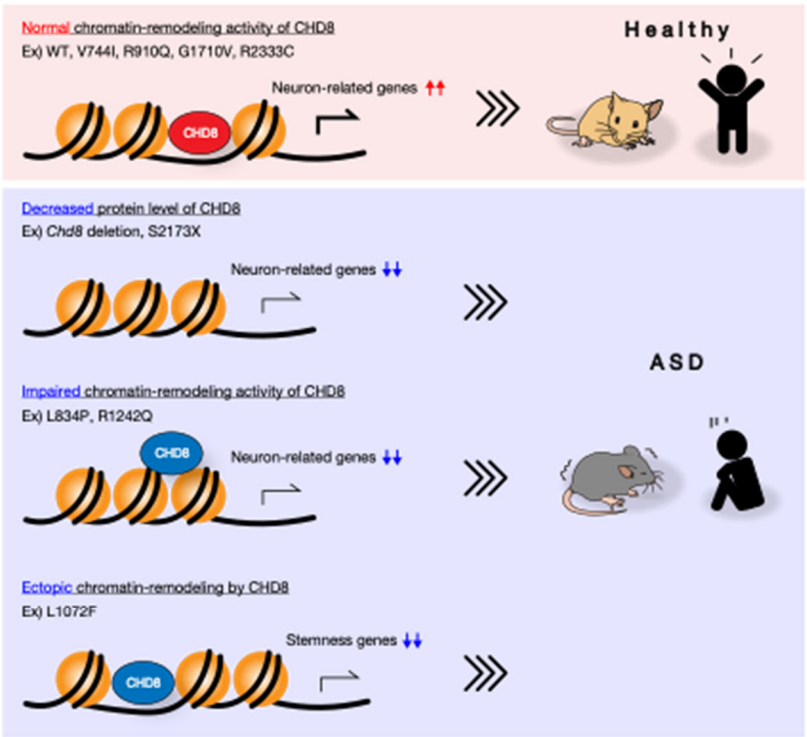
Fig. 6 The L1072F mutation results in ectopic binding of CHD8 to DNA in ESCs and ASD-like behavior in mice. A Total time spent in the light chamber (left) and

number of transitions between the light and dark chambers (right) for adult male mice of the indicated genotypes in the light-dark transition test. **B** Total distance traveled (left) and percentage of time spent in the open arms (right) for adult male mice of the indicated genotypes in the elevated plus-maze test. **C** Total duration of contacts (left) and number of contacts (right) for adult male mice of the indicated genotypes in the social interaction test. Data in **A** to **C** are means ± SEM (n = 20 for each genotype, with the exception that n = 19 for +/+ in the elevated plus-maze test because one mouse fell from the apparatus). *p < 0.05, **p < 0.01, ***p < 0.001 (unpaired two-tailed Student's t test). **D** Volcano plot of RNA-seq data for *Chd8*$^{+/L1072F}$ NPCs compared with WT NPCs at day 14 of induction of differentiation. Upregulated (FDR < 0.05, $\log_2$(fold change) > 0.7) or downregulated (FDR < 0.05, $\log_2$(fold change) < –0.7) genes in mutant versus WT cells are highlighted in red and blue, respectively. The numbers within the plot represent the numbers of differentially expressed genes (DEGs). **E** GO analysis for upregulated or downregulated genes in *Chd8*$^{+/L1072F}$ NPCs at day 14 compared with WT NPCs. **F** Volcano plot of RNA-seq data for *Chd8*$^{+/L1072F}$ ESCs compared with WT ESCs at day 0. Upregulated (FDR < 0.05, $\log_2$(fold change) > 1) or downregulated (FDR < 0.05, $\log_2$(fold change) < –1) genes in mutant versus WT cells are highlighted in red and blue, respectively. The numbers within the plot represent the numbers of differentially expressed genes. **G** Immunoblot analysis of CHD8 in ESCs of the indicated genotypes. **H** Heat map clustering of CHD8 ChIP-seq peaks around the TSS of genes in WT and mutant ESCs. **I** UpSet plot for the number of genes targeted by WT and mutant CHD8 in ESCs. GO analysis for CHD8(L1072F)-specific target genes (1689 genes) is shown in the box. **J** Heat map for the expression levels of representative genes included in the GO terms of stem cell population maintenance and Wnt signaling pathway in WT and *Chd8*$^{+/L1072F}$ ESCs. RNA-seq data for three biological replicates for each genotype are shown.

A

Mutation Functional domain	V744I Chromo	L834P Helicase	R910Q Helicase	L1072F Helicase	R1242Q Helicase	G1710V DNA-bind	S2173X	R2333C BRK
Predicted damage scores	Low	High	Low	High	High	Low	N.A.	Low
Chromatin-remodeling activity	Normal	Impaired	Normal	Normal	Impaired	Normal	Impaired	Normal
Expression level of CHD8 protein	N.A.	N.A.	Normal	Normal	Normal	Normal	Decrease	N.A.
Mouse embryonic lethality for *KI/KI mice*	N.A.	N.A.	Alive	Alive	Lethal	Alive	Lethal	N.A.
Anxiety-like behavior (in the OF, LDT, EPM tests)	N.A.	N.A.	No change	Increase	Increase	No change	Increase	N.A.
Social interaction (duration of contacts in the SI test)	N.A.	N.A.	No change	Increase	Increase	No change	Increase	N.A.
Expression levels of neural differentiation genes	N.A.	N.A.	Normal	Decrease	Decrease	N.A.	N.A.	N.A.

B



Shiraishi *et al.* Figure 7

Fig. 7 Molecular basis for ASD etiology due to missense mutations of the *CHD8* gene. A Summary of the effects of *CHD8* mutations determined in the present study.

N.A., not analyzed; OF, open field; LDT, light-dark transition; EPM, elevated plus-maze; SI, social interaction. **B** Model consistent with our data showing the several mechanisms by which *CHD8* mutations might give rise to the development of ASD.