Role of an atypical cadherin gene, *Cdh23* in prepulse inhibition and its implication in schizophrenia

Shabeesh Balan [1*], Tetsuo Ohnishi [1], Akiko Watanabe [1], Hisako Ohba [1], Yoshimi Iwayama [1], Manabu Toyoshima [1], Tomonori Hara [1], Yasuko Hisano [1], Yuki Miyasaka [2, 3], Tomoko Toyota [1], Chie Shimamoto-Mitsuyama [1], Motoko Maekawa [1, 4], Shusuke Numata [5], Tetsuro Ohmori [5], Tomomi Shimogori [6*], Brian Dean [7, 8], Yoshiaki Kikkawa [2], Takeshi Hayashi [9], Takeo Yoshikawa [1*]

1 Laboratory for Molecular Psychiatry, RIKEN Center for Brain Science, Wako, Saitama, Japan

2 Deafness Project, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Tokyo, Japan

3 Division of Experimental Animals, Graduate School of Medicine, Nagoya University, Nagoya, Japan

4 Department of Biological Science, Graduate School of Humanities and Science, Ochanomizu University, Tokyo, Japan

5 Department of Psychiatry, Institute of Biomedical Science, Tokushima University Graduate School, Tokushima, Japan

6 Laboratory for Molecular Mechanisms of Brain Development, RIKEN Center for Brain Science, Wako, Saitama, Japan

7 The Florey Institute of Neuroscience and Mental Health, Howard Florey Laboratories, The University of Melbourne, Victoria, Australia

8 The Centre for Mental Health, Swinburne University, Victoria, Australia

9 Agricultural Artificial Intelligence (AI) Research Office, Research Center for Agricultural Information Technology, National Agriculture and Food Research Organization (NARO), Tokyo, Japan

*Correspondence:

Shabeesh Balan PhD
Laboratory for Molecular Psychiatry
RIKEN Center for Brain Science
2-1 Hirosawa, Wako, Saitama, 351-0198
Japan
Tel: +81(Japan)-48-462-1111 ext 7634
Fax: +81(Japan)-48-467-7462
E-mail: shabeeshbalan@riken.jp

Tomomi Shimogori PhD
Laboratory for Molecular Mechanisms of Brain Development
RIKEN Center for Brain Science
2-1 Hirosawa, Wako, Saitama 351-0198
Japan
Tel: +81(Japan)-48-467-9779
Fax: +81(Japan)-48-467-7270
E-mail: tomomi.shimogori@riken.jp

Takeo Yoshikawa, MD, PhD
Laboratory for Molecular Psychiatry, RIKEN Center for Brain Science
2-1 Hirosawa, Wako, Saitama 351-0198
Japan
Tel: +81(Japan)-48-467-5968
Fax: +81(Japan)-48-467-7462
E-mail: takeo.yoshikawa@riken.jp

Abstract

We previously identified quantitative trait loci (QTL) for prepulse inhibition (PPI), an endophenotype of schizophrenia, on mouse chromosome 10 and reported *Fabp7* as a candidate gene from an analysis of F2 mice from inbred strains with high (C57BL/6N; B6) and low (C3H/HeN; C3H) PPI levels. The extremely high logarithm of odds (LOD) score of the chromosome-10 QTL suggests the presence of additional causative gene(s). Here, we reanalyzed the previously reported QTLs with increased marker density. The highest LOD score (26.66) peaked at a synonymous coding and splice-site variant, c.753G>A (rs257098870), in the *Cdh23* gene on chromosome 10; the c.753G (C3H) allele showed a PPI-lowering effect. Bayesian multiple QTL mapping also supported the same variant with a posterior probability of 1. Therefore, we examined the roles of *Cdh23* and c.753G>A in PPI. The *Cdh23* gene was shown to be expressed in brain regions relevant for sensorimotor gating. We engineered the c.753G (C3H) allele into the B6 genetic background, which dampened PPI without affecting hearing acuity. An e-QTL (expression-QTL) effect imparted by the c.753G>A variant for the *Cdh23* transcript in the brain was also revealed. The syntenous variant (c.753G>A; rs769896655) in *CDH23* showed a nominally significant enrichment in individuals with schizophrenia. We also identified multiple potentially deleterious *CDH23* variants in individuals with schizophrenia. Collectively, the data in the present study reveal a PPI-regulating *Cdh23* variant and a possible contribution of this gene to schizophrenia susceptibility.

Keywords: prepulse inhibition, quantitative trait locus, *Cdh23*, schizophrenia, hearing loss

Introduction

Deciphering neurobiological correlates for behavioral phenotypes in terms of genetic predispositions that affect molecular, cellular and circuit-level functions is crucial for understanding the pathogenesis of psychiatric disorders [1]. However, polygenicity and the inherent limitations of psychiatric diagnosis based on the subjective experiences of patients have impeded this endeavor [2, 3]. With efforts to classify psychiatric disorders based on the dimensions of observable behavioral and neurobiological measures, endophenotype-based approaches for elucidating genetic liability have attracted interest because these approaches could mitigate clinical heterogeneity [4, 5].

Among the behavioral endophenotypes, the prepulse inhibition (PPI) of acoustic startle response, a reflection of sensorimotor gating, has been consistently reported to be dampened in psychiatric disorders, particularly in schizophrenia [6, 7]. As a robust and heritable endophenotype in schizophrenia [8-11], the genes/variants conferring the risk of dampened PPI can help to elucidate the genetic architecture of schizophrenia [12, 13]. We have previously performed a large-scale quantitative trait loci (QTL) analysis for PPI and mapped six major loci through an analysis of 1,010 F2 mice derived by crossing selected inbred mouse strains with high (C57BL/6NCrlCrlj: B6Nj) and low (C3H/HeNCrlCrlj: C3HNj) PPI [14]. Among these six loci, the chromosome-10 QTL showed the highest logarithm of odds (LOD) score, and we identified fatty acid-binding protein 7 (*Fabp7*) as one of the candidate genes in the QTL for regulating PPI and schizophrenia pathogenesis [14, 15]. However, this high LOD score cannot be explained by a single gene, considering the low phenotypic variance attributed to the individual genes/markers, suggesting the presence of additional causative gene(s) [16]. Therefore, in the current study, we aimed to (i) delineate additional gene(s) that regulate the PPI phenotype by reanalyzing the QTL with higher marker density, (ii) validate the candidate gene variant(s) by analyzing causal allele knock-in

mice, and (iii) examine the potential role of candidate gene in the predisposition to schizophrenia.

Results

Dense mapping of QTL reveals a high LOD score on chromosome 10 and *Cdh23* as a putative candidate gene

We reanalyzed the QTL for the PPI levels phenotyped in 1,010 F2 mice derived from F1 parents (B6NjC3HNj) from our previous study (Supplementary Figure 1A, B) [14]. Here we increased the density of the markers, in addition to the previously genotyped microsatellites, for the QTL analysis. To this end, exome sequencing of C57BL/6NCrl (B6N) and C3H/HeNCrl (C3HN) mouse strains was performed (data not shown) and 148 strain-specific single nucleotide variants (SNVs) from the six previously reported loci on chr1, chr3, chr7, chr10, chr11 and chr13 were additionally genotyped (Supplementary Table 1). Note here that the previously used mouse strains C57BL/6NCrlCrlj (B6Nj) and C3H/HeNCrlCrlj (C3HNj) are no longer available because their supply ended in 2014 [16]. Composite interval mapping (CIM) consistently revealed QTL signals with high LOD scores on chromosome 10 for different prepulse levels (Supplementary Figure 1C). The PPI at a prepulse level of 86 dB showed the highest LOD score, 26.66, which peaked at a synonymous coding and splice-site variant c.753G>A (rs257098870) in the cadherin 23 (*Cdh23*) gene (Figure 1A, B). This QTL accounted for 11.7% of the phenotypic variance (Supplementary Table 2). Additionally, Bayesian multiple QTL mapping performed for PPI at a prepulse level of 86 dB also uncovered the same variant as the QTL with a posterior probability of 1, though accounting for less phenotypic variance than was suggested by CIM (Figure 1C, Supplementary Table 3). This lower variance could be attributed to the effects of multiple QTLs (markers) simultaneously fitted in the model. Notably, the chromosome-10 QTL encompassing *Cdh23*

was more robust among males when dichotomized for gender in both the CIM and Bayesian methods (Supplementary Table 2, 3, Supplementary Figure 1D). Thus, we set out to pursue the role of *Cdh23* as a putative candidate gene.

Cdh23 is predominantly expressed in the subthalamic and pontine regions of the brain

Cdh23, an atypical cadherin, is a member of the cadherin superfamily, which encodes calcium-dependent cell-cell adhesion glycoproteins [17, 18]. The role of *Cdh23* in cochlear hair cell architecture and hearing has been extensively studied [19]. Although many members of the cadherin superfamily are expressed in neural cells and implicated in neurodevelopment, evidence for the role of *Cdh23* in neuronal functions is minimal. Previous studies have shown *Cdh23* expression in the brains of zebrafish [20], and in the auditory cortex [21, 22] and the zona incerta of mouse brain [23]. To query the potential role of *Cdh23* in neuronal function and, in turn PPI, we evaluated its expression in the mouse brain. RNA *in situ* hybridization showed that *Cdh23* is predominantly expressed in subthalamic (including zona incerta and subthalamic nucleus) and pontine regions in both B6N and C3HN mouse strains (Figure 1D) across the developmental stages (Supplementary Figure 2). There were no distinct differences in expression patterns between the genders (Supplementary Figure 3A). Furthermore, a similar expression pattern was observed in the brain regions from the marmoset (*Callithrix jacchus*) (Supplementary Figure 3B), indicating functional conservation across species. Previous reports have shown the potential role of these brain regions in sensorimotor gating [24]. In humans, *CDH23* transcript expression was shown to be developmentally regulated across different brain regions, which was also observed in the *in vitro* neuronal differentiation lineage of human induced pluripotent stem cells (hiPSCs) (Figure 1E, F) [25, 26]. These lines of evidence indicate potential roles for Cdh23 in neuronal function and PPI.

Cdh23 c.753G>A is the only distinct coding variant between B6N and C3HN mouse strains

Since a high LOD score was found for the SNV marker c.753G>A (rs257098870) in *Cdh23*, which differed between the B6N and C3HN strains, we reasoned that this variant or other flanking functional variants in the *Cdh23* gene might be responsible for the observed phenotypic effect. To identify the putative causal genetic variant(s), all the coding and flanking intronic regions of the *Cdh23* gene in B6N and C3HN mice were resequenced. However, no other variants were identified, except for the same genotyped synonymous coding and splice site variant, c.753G>A (rs257098870), on the 3' end of exon 9 with the G allele in C3HN mice and the A allele in B6N mice (Supplementary Figure 4 A-D). The c.753A allele was previously known to cause age-related hearing loss (ARHL) by affecting cochlear stereocilia architecture [27-29]. Although the variant does not cause an amino acid change (P251P), c.753A results in skipping of exon 9 [27]. We also queried *Cdh23* genetic variations in C57BL/6NJ and C3H/HeJ mouse strains from Mouse Genomes Project data (https://www.sanger.ac.uk/science/data/mouse-genomes-project/), which again showed c.753G>A as the only distinct coding variant between the strains (Supplementary Table 4). Variant allele A was specific for some of the inbred mouse strains, and the overall variant allele frequency was relatively high (39%), as observed from the Mouse Genomes Project data (Supplementary Figure 4E, F). Moreover, the presence of the G allele and low PPI levels in inbred mouse strains were observed to be correlated (Supplementary Figure 4G) [14]. Therefore, we reasoned that c.753G might contribute to the dampened PPI, and we set out to evaluate its role in PPI function.

C3HN strain-specific *Cdh23* c.753G allele knock-in mice showed reduced PPI levels without affecting hearing acuity

To test the role of the *Cdh23* c.753G>A variant in modulating PPI, we generated a mouse model in the B6N genetic background in which the C3HN strain-specific *Cdh23* c.753G allele was knocked in using CRISPR/Cas9n-mediated genome engineering (Supplementary Figure 5A-C). Mice homozygous for the knock-in allele (GG) and control littermates with the wild-type allele (AA) derived from heterozygous mice were used for the experiments. Since the A allele in the B6N strain results in the skipping of exon 9 and also causes ARHL, knock-in of the G allele in the B6 genetic background should reverse these events, and lead to dampening of the PPI.

We tested the splicing pattern of *Cdh23* exon 9 in the brain regions (subthalamic and pontine regions) where *Cdh23* was predominantly expressed from knock-in mice homozygous for the G allele, control littermates homozygous for the A allele, and inbred B6N and C3HN mice. Exon 9 was almost completely retained in G allele-bearing mice in comparison to the A allele, which showed partial skipping of exon 9 (Figure 2A). This was in line with the alternative splicing observed in inbred B6N and C3HN mouse brains (Figure 2A). However, the splicing pattern was tissue-specific, such that the expression of *Cdh23* transcript lacking exon 9 was lower in the pontine and subthalamic regions, than in the inner ear (Supplementary Figure 6A). The almost exclusive expression of exon 9-containing transcripts in the inner ear from the *Cdh23* c.753G allele knock-in mouse was in agreement with reports from inbred strains with the G allele, thus ensuring successful knock-in (Supplementary Figure 6A)[27] .

Next, we tested the PPI levels in *Cdh23* c.753G allele knock-in mice and *Cdh23* c.753A allele litter mates (13 weeks old). The G allele knock-in mice showed lower PPI

levels than the A allele mice in all the prepulse levels tested (Figure 2B). Dampened PPI was observed in both male and female mice (Supplementary Figure 6B). Since the *Cdh23* c.753A allele is reported in ARHL, we further tested the hearing acuities of knock-in mice (GG) and control littermates (AA) to rule out any confounding effect. To this end, we examined the auditory brainstem response (ABR) thresholds for tone-pip stimuli at 8, 16 and 24 kHz in 13-week-old mice (the age at which PPI was measured) for both genotypes. We found no differences in the ABR threshold between AA and GG genotypes (Figure 2C), thus supporting the conclusion that the observed impairment of PPI in the *Cdh23* c.753G allele knock-in mouse was not confounded by hearing acuity, and that the *Cdh23* c.753A allele did not elicit hearing impairment at least up to 13 weeks of age. We also examined ABR threshold of the knock-in mice at 6 months, the age at which ARHL is known to manifest [28, 29]. We observed a significantly higher ABR threshold in the *Cdh23* c.753A allele knock-in mouse than in the G allele, indicating the onset of ARHL in the AA genotype at 6 months (Figure 2D). These observations were further corroborated with the evidence from scanning electron microscopy (SEM) images of cochlear hair cells (middle area), where the stereociliary architecture was intact in both genotypes at 8 weeks (Figure 2E). The stereociliary morphology was preserved in 6-month-old *Cdh23* c.753G allele knock-in mice with the B6N genetic background when compared to the *Cdh23* c.753A allele littermates, which showed loss of hair cells, indicating ARHL (Figure 2E, Supplementary Figure 7A). These results were in agreement with recent studies that showed the rescue of hearing acuity and stereocilia degeneration, upon G allele knock-in in different models of hearing loss resulting from the *Cdh23* c.753G-to-A mutation [29-31]. Taken together, these findings revealed that the C3HN strain-specific *Cdh23* c.753G is the causal allele that dampens PPI levels and that this effect is not confounded by hearing acuity.

Potential role of the *Cdh23* c.753G>A variant in *Cdh23* transcript expression

To evaluate the functional role of the *Cdh23* c.753G>A variant, we examined the transcript expression levels of *Cdh23* in pontine and subthalamic regions of *Cdh23* c.753G allele knock-in mice in comparison to the *Cdh23* c.753A allele littermates using digital PCR. We observed significantly lower *Cdh23* transcript levels in the G allele carriers in both brain regions, indicating that c.753G>A acts as an expression QTL (e-QTL) (Figure 3A). However, this trend was not observed in inbred adult B6N and C3HN mice (4-6 weeks-old) (Figure 3B). Interestingly, differences in *Cdh23* transcript expression, between inbred B6N and C3HN mice, was observed during earlier stages of development particularly in E16.5 and P0 (Figure 3C). Considering the developmentally regulated expression pattern of *Cdh23*, these results suggest that the e-QTL variant might act differentially during the early neurodevelopmental stages. Furthermore, it also points to the presence of epistatic interaction between the c.753 locus and genetic backgrounds for regulating the *Cdh23* transcript expression.

Recent evidence has suggested that dysfunction of GABAergic neurons is elicited during neurodevelopment by *Cdh23* deficiency [21], and the reduction of GABAergic neurons has been implicated in the PPI deficits in mice [32]. To test this, we evaluated the transcript expression levels of *Pvalb*, *Gad1* and *Slc32a1*, which are coexpressed with *Cdh23* in subthalamic region [33], from knock-in mice and control littermates. However, we did not observe any significant differences in the transcript expression levels of these genes (Supplementary Figure 7B), which might be attributable to bulk sampling from the distinct molecular domains of subthalamic region with heterogeneous cell types [34].

Role of *CDH23* in schizophrenia

Schizophrenia-like psychosis was reported in Usher syndrome caused by mutations in *CDH23* [35-37]. Furthermore, a recent genome-wide association study (GWAS) of PPI in schizophrenia also revealed association signals for *CDH23* [13]. We therefore examined whether novel rare loss-of-function (LoF) variants in the coding regions of *CDH23* are enriched in Japanese schizophrenia cases, using molecular inversion probe (MIP)-based sequencing method (Supplementary Figure 8). Several patient-specific novel variants with potential deleterious effects predicted by *in silico* tools were identified in *CDH23*. But each variant was very rare and was observed in 1-2 cases only (Figure 3D, Supplementary Figure 9A). In addition, schizophrenia exome meta-analysis consortium data revealed no enrichment of *CDH23* LoF variants in schizophrenia (Supplementary Figure 9B). The constraint matrices from the Genome Aggregation Database (gnomAD) revealed that *CDH23* is tolerant to mutations (Supplementary Figure 9C).

Interestingly, a syntenous variant for *Cdh23* c.753G>A was observed in humans (rs769896655; c.753G>A; P251P), with G as the major allele (Figure 3D, E, F). The variant allele A was more frequent in the Japanese population than others in the gnomAD database [38] (Supplementary Table 5). When we tested the enrichment of this variant in schizophrenia, a nominally significant over representation of A-allele in Japanese population was observed (schizophrenia vs. controls; $P = 0.045$) (Table 1), particularly in the Honshu region (the main island of Japan). However, the enrichment did not persist when the samples from Honshu and Shikoku regions were analysed together ($P = 0.27$) (Supplementary Table 6). This could be due to the population genetic structure of the Shikoku region, which is distinct from the main island of Honshu, and is genetically closer to Han Chinese [39]. The pooled analysis of schizophrenia samples [Japanese + Schizophrenia exome meta-analysis consortium, (SCHEMA); https://schema.broadinstitute.org/] and all controls, also did not

show any allelic association ($P = 0.21$) (Supplementary Table 6), which could be attributed to the rarity of variant allele A in other populations compared to the Japanese population. Remarkably, the PPI-associated variants in the *CDH23* locus from the GWAS are in considerable linkage disequilibrium with rs769896655, indicating the potential role of *CDH23* in schizophrenia (Supplementary Figure 10A) [13]. *CDH23* transcript expression in postmortem brain samples from Brodmann area 8 was not significantly different between individuals with schizophrenia and controls (Figure 3G). Allele-specific expression of the variant rs769896655 in the *CDH23* transcript was tested in hair follicles, peripheral blood samples, hiPSCs, hiPSC-derived neurospheres, and neurons from a healthy subject heterozygous for the variant (Supplementary Figure 10B). A statistically significant preferential inclusion of the G allele was observed in *CDH23* transcripts from peripheral blood samples and hiPSCs (Figure 3H), indicating that allele-specific expression is cell/tissue-specific and developmental stage-dependent.

Discussion

By reanalyzing and fine-mapping mouse chromosome-10 PPI-QTL [14], we revealed *Cdh23* as a candidate for the PPI endophenotype. Our prior studies hinted at a possible role of *Cdh23*, but did not address this directly, as the gene is in the ARHL locus [14, 16]. We showed here that the C3H/HeNCrl (C3HN) strain-specific G allele of the *Cdh23* c.753G>A variant is causal for low PPI; we engineered the c.735G into the C57BL/6NCrl (B6N) genetic background, which dampened the PPI without affecting hearing acuity, as evidenced by the ABR threshold or stereocilia architecture. We also observed a potential e-QTL effect of the G allele in lowering *Cdh23* transcript expression in subthalamic and pontine regions.

Cdh23, an atypical cadherin, is a member of the cadherin superfamily which includes multiple subtypes involved in calcium-dependent cell–cell adhesion and/or signaling events

through homotypic and heterotypic interactions [18]. Although several members of the cadherin superfamily govern multiple facets of neurodevelopment and function, evidence for atypical cadherins is limited [17]. The first evidence for the role of Cdh23 in neurodevelopment showed a lower number of parvalbumin-positive interneurons in the auditory cortex of *Cdh2*3-deficient mice, resulting from deficits in motility and/or polarity of migrating interneuron precursors during early embryonic development [21]. Recently, *Cdh23* was also shown to be involved in mouse brain morphogenesis [40]. Impaired PPI has been attributed to neurodevelopmental deficits [41]. Notably, the expression of *Cdh23* in the subthalamic and pontine regions is conserved across the species, as shown in mouse and marmoset in this study, and those regions form the neural circuitry that regulates PPI [24, 42-44]. Interestingly, deep brain stimulation (DBS) in the subthalamic region has been shown to improve PPI in rodent models and to ameliorate psychosis symptoms in individuals with Parkinson's disease [45, 46].

Moreover, mounting evidence from genetic studies has indicated the potential role of *CDH23* in schizophrenia and other neuropsychiatric disorders [47, 48]. Although we could not show any enrichment of *CDH23* variants in schizophrenia, a syntenous variant of mouse *Cdh23* c.753G>A in humans (rs769896655) showed a nominally significant enrichment, with overrepresentation of the A allele in Japanese population. The variant allele is rare in other populations, showing population specificity. This could be a reason for the lack of genetic association in the pooled analysis with SCHEMA data, which consists predominantly of European samples. Interestingly, a recent GWAS for the PPI phenotype in individuals with schizophrenia revealed a potential role of *CDH23* [13]. Although not genome-wide significant, the lead variant in the *CDH23* interval is in considerable linkage disequilibrium with rs769896655, thus supporting *CDH23* as a candidate schizophrenia susceptibility gene.

Although the c.753-G allele was causal in lowering PPI in mice, the reason for the overrepresentation of A allele in schizophrenia cases is elusive. The wild mice sampled across different geographical locations showed the G allele in a homozygous state [49]. It may be interesting to see whether low PPI is evolutionarily advantageous for mice to survive in the wild. A low PPI may help to enhance vigilance. In mice, the variant allele A arose in specific strains, including B6N, when wild mice were inbred to develop strains with a homogenous genetic background, and the variant allele may have primed the variant-associated phenotypes [50, 51]. The rarity of the variant A allele in humans, the lack of homozygotes and the population specificity indicate that the syntenous variant might have originated recently. It is possible that CDH23/Cdh23 may have a potential role in the PPI phenotype, but regulation of the phenotype through the syntenous variant might not be conserved. The species specificity of variants in the manifestation of phenotypes is known in Parkinson's disorder, where the disease-causing variant of alpha-synuclein in humans is commonly observed in rodents and other model organisms that do not display any pathological features [52]. Regarding rs769896655 in *CDH23*, the rarity of variant allele A has limited the testing of genotype-phenotype (PPI/hearing loss) correlations in humans. Future studies to test *CDH23* expression in postmortem brain samples from schizophrenia cases and controls in subthalamic and pontine regions are warranted.

In summary, the current study demonstrates the role of an atypical cadherin, Cdh23, in regulating PPI through a genetic variant without affecting hearing acuity. Future studies are required to precisely delineate the role of the *Cdh23* genetic variant in neuronal function and PPI regulation. Additionally, to translate the findings to humans, large-scale studies to test the association of *CDH23* c.753G>A in schizophrenia cases are warranted in the Japanese population, where the variant allele is relatively prevalent.

Materials and Methods

Study approval

All animal experiments were performed in compliance with relevant laws and guidelines, and were approved by the Animal Ethics Committee at RIKEN. Studies involving human subjects were approved by the Human Ethics Committee at RIKEN. All participants in the genetic studies gave informed, written consent to join the study after receiving a full explanation of the study protocols and objectives. All postmortem brain study procedures were carried out with informed written consent from the next of kin, procured by the brain bank.

Animals

For the QTL analysis, we genotyped F2 mice from our previous study [14] produced by randomly intercrossing the F1 hybrid mice resulting from a cross of male C3H/HeNCrlCrlj (C3HNj) and female C57BL/6NCrlCrlj (B6Nj) mice obtained from Japan's Charles River Laboratories (Yokohama, Japan). For other experiments, C57BL/6NCrl (B6N) and C3H/HeNCrl (C3HN) mice were used.

Human DNA samples

For resequencing all the protein-coding exons of the CDH23 gene, a total of 1,200 individuals with schizophrenia (diagnosed according to DSM-IV) of Japanese descent (657 men, mean age 49.1±14.0 years; 543 women, mean age 51.1±14.4 years) were used. The diagnosis of schizophrenia was made using the Diagnosis and Statistical Manual of Mental Disorders IV (DSM-IV) criteria and was confirmed by at least two experienced psychiatrists. For subsequent genotyping, an additional 811 individuals with schizophrenia were included, for a total of 2,011 affected individuals (1,111 men, mean age 47.2 ± 14.1 years; 901 women, mean age 49.2 ± 14.7 years), along with 2,170 healthy controls (889 men, mean age 39.2

±13.8 years; 1,281 women, mean age 44.6 ± 14.1 years). The samples were recruited from the Honshu region of Japan (the main island), where the population fall into a single genetic cluster [53]. Additionally, we also used genetic data of 8,380 healthy Japanese controls from the Integrative Japanese Genome Variation Database (iJGVD) by Tohoku Medical Megabank Organization (ToMMo) (https://ijgvd.megabank.tohoku.ac.jp/) to compare the allele frequencies of the identified variants. The variant rs769896655 was genotyped in another sample set from Shikoku region (one of the islands in Japan), which included 1,062 individuals affected with schizophrenia (602 men, mean age 56.37 ± 13.59 years; 460 women, mean age 59.53 ± 14.84 years), and 1,974 healthy controls (798 men, mean age 35.32 ±13.8 years; 1,176 women, mean age 41.98 ± 12.79 years).

Post mortem brain tissues

Tissues from Brodmann area 8 (BA8) from individuals with schizophrenia (n = 50) and controls (n = 71) were obtained from the Victorian Brain Bank Network, Florey Institute for Neuroscience and Mental Health [54-56]. The diagnosis of schizophrenia was made after reviewing the clinical records, based on DSM-IV criteria and the Diagnostic Instrument for Brain Studies, which allows a consensus psychiatric diagnosis to be made after death [57, 58]. To exclude any psychiatric or neurological disorders in the controls, the case histories were extensively reviewed, along with the interview of the treating clinicians and family members.

Human induced pluripotent stem cells (hiPSCs)

To test the regulation of *CDH23* expression during neurodevelopment *in vitro*, hiPSC lines established from healthy controls (n = 4) as reported previously were used [25]. To analyze allele-specific expression of *CDH23* for the rs769896655 (c.753 G>A) variant, another hiPSC line was established from a healthy individual (TKUR120) who was heterozygous for

the variant. Establishment of hiPSCs and their differentiation to neurospheres and neurons were performed as described before [26].

Selection of additional single nucleotide variant (SNV) markers for fine mapping of QTL

To perform fine mapping, additional SNV markers were selected from the six previously identified loci (chr1, chr3, chr7, chr10, chr11, chr13) for PPI [14]. To this end, exome sequencing of inbred C3HN and B6N mouse strains was performed as described in the supplementary methods. SNVs were selected from the regions ± 20 Mbp from the central position of each previously genotyped Sequence-Tagged Sites (STS) marker. One SNV marker from each megabase was primarily chosen using the following criteria: 1) specifically detected in either the B6N or C3HN strain, 2) homozygous, 3) with a PASS flag and 4) nearest to the central position of each megabase. If there were SNVs reported in dbSNP 128 within ± 0.25 Mbp of the central position of each megabase, we selected a marker from SNVs in dbSNP (that might not be the SNV nearest to the central position). The regions with no SNVs satisfying these criteria were skipped. If the distance between two markers was less than 0.25 Mbp, one marker was excluded. The list of the 148 selected SNV markers is shown in Supplementary Table 1.

QTL analysis

Reanalysis of PPI-QTL was performed by increasing the density of the analyzed markers. The 148 additionally selected SNV markers were genotyped in 1,012 F2 mice by Illumina BeadArray genotyping (Illumina Golden Gate assay) on the BeadXpress platform as per the manufacturer's instructions. First, we applied a composite interval mapping method [59] for 86 dB-PPI to scan relevant QTLs using Windows QTL Cartographer v2.5 (https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm) on autosomal chromosomes only,

as no QTL signal was detected on the X chromosome in the previous analyses [14]. In the present analysis, the gender of the mice was taken into consideration. The results were expressed as LOD scores, and a threshold of LOD score corresponding to a genome-wide significance level of 1% was determined to be 4.16 with 1000 repetitions of the permutation test. Subsequently, Bayesian multiple QTL mapping was conducted for fine mapping of QTLs at the 86 dB prepulse level. In this Bayesian method, we located a putative QTL on each marker and estimated the posterior probability that the QTL had a nonzero significant effect. The phenotype of the i^{th} F2 mouse, y_i, was written as the following statistical model:

$$y_i = \mu + s_i b + \sum_{l=1}^{N} (\gamma_l u_{il} a_l + \eta_l v_{il} d_l) + e_i$$

where μ was the intercept of the model, b was the gender effect, s_i was a covariate indicating the gender of the ith mouse with s_i =0 and 1 corresponding to male and female, respectively, N was the number of markers on which putative QTLs were located, u_{il} and v_{il} were the covariates indicating the genotype of the ith mouse at the lth marker, with u_{il}=1, 0, -1 and v_{il} = 0, 1, 0 for homozygous for the B6 type and heterozygous and alternative homozygous for the C3 type, respectively, a_l and d_l were additive and dominance effects of the lth marker, γ_l and η_l were indicator variables with γ_l = 1 or 0 meaning the inclusion or exclusion of the additive effect at the QTL in the model fitting, respectively, and analogously, η_l =1 or 0 for inclusion or exclusion of dominance effect, and e_i was the residual error assumed to follow a normal distribution with mean 0 and variance σ_e^2.

In a Bayesian framework, these parameters and indicator variables were evaluated based on their posterior distributions constructed from the prior distributions and the distribution of observed samples, including y_i, u_{il} and v_{il} (i=1,2,…,n; l=1,2,…,N). The posterior distributions can be empirically established by sampling values of the parameters

and indicator variables via the Markov chain Monte Carlo (MCMC) sampling procedure, which is time-consuming. Instead, we applied a variational approximation method to effectively obtain the posterior expectations of the parameters and indicator variables [60]. This variational Bayes method was originally utilized in the context of genomic selection [61], where a model predicting breeding values from SNP genotypes, which is analogous to the model considered here, is constructed to enable individuals with superior genetic performance to be effectively selected based on their SNP genotypes in animal and plant breeding [60]. The existence of QTLs with significant effects was judged using the posterior expectations of indicator variables, γ_l and η_l, for additive and dominance effects, which were regarded as the posterior probabilities of the putative QTLs being fitted in the model. As a result of this variational Bayes analysis, we found no QTLs with a significant dominance effect; thus, only the plot of the posterior expectations of γ_l is shown in Figure 1C.

Phylogenetic and conservation analysis of *Cdh23* c.753G>A

We compared the conservation of the *Cdh23* c.753G>A variant (rs257098870) and the flanking region among eutherian mammals and rodents as described in the supplementary methods.

Generation of *Cdh23* c.753G allele knock-in mice by CRISPR/Cas9n-mediated genome editing

The *Cdh23* c.753G allele knock-in mice in the B6 genetic background were generated by genome editing using the CRISPR/Cas9 nickase [62-64]. Briefly, B6N zygotes obtained by *in vitro* fertilization were microinjected with the cocktail constituting 5 ng/mL Cas9 nickase mRNA (System Biosciences, Mountain View, CA), 5 ng/mL each of two short guide RNAs (sgRNAs) (Supplementary Table 7), which were synthesized *in vitro* (T7 gRNA Smart Nuclease Synthesis Kit, System Biosciences) according to the manufacturer's instructions,

and 5 ng/mL single-stranded oligodeoxynucleotide (ssODN) (Supplementary Table 7). Injected zygotes were transplanted into the uteruses of pseudopregnant dams, and the resulting pups were obtained by cesarean section. The target region of the *Cdh23* gene was directly sequenced from the PCR products amplified from the template DNAs extracted from the tails of the pups. Additionally, the PCR products were subcloned and sequenced in founders with knocked-in alleles. Promising founders were selected, and upon their reaching sexual maturity, *in vitro* fertilization was performed with B6N strain-derived oocytes to check germline transmission and to obtain mice with heterozygous mutated alleles. Furthermore, the heterozygous mice were intercrossed to produce homozygote knock-in and wild-type control littermates for further experiments. Routine genotyping by Sanger sequencing was performed to verify the allele knock-in.

Behavior analysis

The PPI test was performed according to previously published methods [14, 15].

Analysis of auditory brainstem response

Auditory brainstem response was recorded as described previously [65]. Briefly, mice were anesthetized by intraperitoneal injection of sodium pentobarbital (Somunopentil; 60-80 mg/kg) diluted in saline and were placed on a heating pad to maintain body temperature at 37°C during ABR evaluation. Needle electrodes were placed subcutaneously into the vertex (reference), right ear (active), and left ear (ground). The tone stimulus was produced by a speaker (ES1 spc; BioResearch Center Nagoya, Japan) probe inserted into the auditory canal of the right ear. The ABR thresholds from right ears in 13-week-old and 6-month-old mice (*Cdh23* c.753G>A; AA vs GG genotypes) were measured with tone-pip stimuli at 8, 16 and 24 kHz generated by System3 (TDT). The resulting ABR waveforms were bandpass-filtered (< 3 kHz and > 100 kHz), amplified 1,000 times (AC PreAmplifier, P-55, Astro-Med Inc.)

and recorded using the PowerLab 2/25 System (AD Instruments) for 10 microseconds. A total of 500 recordings were averaged. The waveform data were analyzed by LabChart v.7 (AD Instruments). The ABR thresholds were obtained for each stimulus by reducing the sound pressure level (SPL) for 10 dB steps (80 dB to 10 dB) to identify the lowest level at which an ABR pattern could be reliably detected. The ABR threshold was confirmed for consensus by independent evaluation of three investigators blinded to the genotype of the tested mice.

Cochlear hair cell architecture analysis

Scanning electron microscopy (SEM) of cochlear stereocilia was performed as described previously [29, 66] in four ears from *Cdh23* c.753G allele knock-in mice and *Cdh23* c.753A (13 weeks and 6 months of age) and examined using a Hitachi S-4800 field emission scanning electron microscope at an accelerating voltage of 10 kV.

Gene expression analysis

To detect the localization of *Cdh23* expression in the brain, we performed RNA *in situ* hybridization in mouse (C3 and B6) (postnatal days: P0, P6, P14, 3weeks, 4 weeks, 5 weeks, 6 weeks, and 8 weeks-old) and marmoset (neonate and 3 months-old) as described previously (https://gene-atlas.brainminds.riken.jp/) [67, 68]. We also queried the expression of *CDH23* in the human developmental brain transcriptome (http://development.psychencode.org/).

Cdh23/CDH23 mRNA expression was queried in (i) mouse brain samples, viz., subthalamic region/zona incerta and pontine region; (ii) hair follicle, blood, hiPSC, hiPSC-derived neurospheres, and neurons established from the control sample; and (iii) postmortem brain samples (BA8) from individuals with schizophrenia and controls. Total RNA was extracted using the miRNeasy Mini kit (QIAGEN GmbH, Hilden, Germany), and single-stranded cDNA was synthesized using the SuperScript VILO cDNA synthesis kit according

to the manufacturer's instructions. Splicing and exon skipping resulting from the *Cdh23* c.753G>A variant were tested by semiquantitative RT-PCR normalized to *Gapdh* (Supplementary Table 8). The mRNA levels for the target genes (Supplementary Table 8) were quantified by real-time quantitative RT-PCR using TaqMan Gene Expression Assays, performed in triplicate, based on the standard curve method, normalized to GAPDH/Gapdh as an internal control. To test the role of the *Cdh23* c.753G>A variant as a cis expression QTL (e-QTL) for *Cdh23* expression, digital PCR was performed in mouse brain samples from the subthalamic region/zona incerta and pontine region using standard procedures and TaqMan Gene Expression Assays in a QuantStudio™ 3D Digital PCR System (Life Technologies Co., Carlsbad, CA, USA). Values outside the mean ± 2 SD in the group were considered outliers and were omitted from the analyses.

Mutation analysis of *CDH23* in schizophrenia by targeted next-generation sequencing (NGS) using molecular inversion probes

Genomic DNA was isolated from blood samples obtained from human subjects using standard methods. MIPs were designed for the *CDH23* gene using MIPgen [69] (http://shendurelab.github.io/MIPGEN/), targeting the coding exons and the flanking exon-intron boundaries of the genes (GRCh37 build) and covering all transcripts (Supplementary Table 9). A total of 152 MIPs were designed for *CDH23*, aligning to the design parameters as reported previously [70]. Library preparation, sequencing and variant analysis were performed as described in the supplementary methods.

Statistical analysis

The data in the figures are represented as the mean ± SEM. Statistical analysis and graphical visualization was performed using GraphPad Prism 6 (GraphPad Software). The total sample size (*n*) and the analysis methods are described in the respective figure legends. Statistical

significance was determined using a two-tailed Student's *t*-test and the Holm-Sidak method. When multiple comparisons were needed, two-way ANOVA with Bonferroni's correction was performed. Differences in the gene expression levels in postmortem brain samples between the cases and controls were detected by the Mann–Whitney *U*-test (two-tailed). A *P* value of <0.05 was considered statistically significant.

Acknowledgements

We are grateful to the Animal Resources Development and Support Units for Bio-Material Analysis at RIKEN CBS Research Resources Division for animal maintenance, embryo manipulation, and DNA sequencing services.

Funding

This work was supported by the Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellowship for Overseas Researchers (grant number 16F16419) [S.B], and Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology (grant number JP19H05435) [T.Y.]. The funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

The authors declare no competing interests

Author contributions

Conceptualization: TY, SB, TO; Data curation: SB, TO, AW, HO, YI; Formal analysis: SB, TH, TO, HO, AW, YI, YK; Funding acquisition: TY, SB; Investigation: SB, TO, TH, AW, HO, YI, MT, ToH, YH, YM, SN, TS, YK; Methodology: SB, TO, YH, TS, TH; Project administration: TY, SB, TO; Resources: TY, TT, MT, MM, BD, YK, TOhm, TS, CSM, TH; Software: TH, TS; Supervision: TY, SB, TO, TS; Validation: SB, TO, SN, YI; Visualization: SB, YK, YM; Roles/Writing - original draft: SB; Writing - review & editing: SB, TY, TO, TH, YK

References

1. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. The American journal of psychiatry. 2010;167(7):748-51. doi: 10.1176/appi.ajp.2010.09091379. PubMed PMID: 20595427.

2. Smoller JW, Andreassen OA, Edenberg HJ, Faraone SV, Glatt SJ, Kendler KS. Psychiatric genetics and the structure of psychopathology. Mol Psychiatry. 2019;24(3):409-20. doi: 10.1038/s41380-017-0010-4. PubMed PMID: 29317742; PubMed Central PMCID: PMCPMC6684352.

3. Demkow U, Wolanczyk T. Genetic tests in major psychiatric disorders-integrating molecular medicine with clinical psychiatry-why is it so difficult? Transl Psychiatry. 2017;7(6):e1151. doi: 10.1038/tp.2017.106. PubMed PMID: 28608853; PubMed Central PMCID: PMCPMC5537634.

4. Gottesman, II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. The American journal of psychiatry. 2003;160(4):636-45. doi: 10.1176/appi.ajp.160.4.636. PubMed PMID: 12668349.

5. Sanchez-Roige S, Palmer AA. Emerging phenotyping strategies will advance our understanding of psychiatric genetics. Nat Neurosci. 2020;23(4):475-80. doi: 10.1038/s41593-020-0609-7. PubMed PMID: 32231337.

6. Braff DL, Geyer MA, Swerdlow NR. Human studies of prepulse inhibition of startle: normal subjects, patient groups, and pharmacological studies. Psychopharmacology (Berl). 2001;156(2-3):234-58. doi: 10.1007/s002130100810. PubMed PMID: 11549226.

7. Swerdlow NR, Braff DL, Geyer MA. Sensorimotor gating of the startle reflex: what we said 25 years ago, what has happened since then, and what comes next. J Psychopharmacol. 2016;30(11):1072-81. doi: 10.1177/0269881116661075. PubMed PMID: 27539931; PubMed Central PMCID: PMCPMC6036900.

8. Greenwood TA, Braff DL, Light GA, Cadenhead KS, Calkins ME, Dobie DJ, et al. Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. Arch Gen Psychiatry. 2007;64(11):1242-50. doi: 10.1001/archpsyc.64.11.1242. PubMed PMID: 17984393.

9. Seidman LJ, Hellemann G, Nuechterlein KH, Greenwood TA, Braff DL, Cadenhead KS, et al. Factor structure and heritability of endophenotypes in schizophrenia: findings from the Consortium on the Genetics of Schizophrenia (COGS-1). Schizophr Res. 2015;163(1-3):73-9. doi: 10.1016/j.schres.2015.01.027. PubMed PMID: 25682549; PubMed Central PMCID: PMCPMC5944296.

10. Swerdlow NR, Light GA, Sprock J, Calkins ME, Green MF, Greenwood TA, et al. Deficient prepulse inhibition in schizophrenia detected by the multi-site COGS. Schizophr Res. 2014;152(2-3):503-12. doi: 10.1016/j.schres.2013.12.004. PubMed PMID: 24405980; PubMed Central PMCID: PMCPMC3960985.

11. Swerdlow NR, Light GA, Thomas ML, Sprock J, Calkins ME, Green MF, et al. Deficient prepulse inhibition in schizophrenia in a multi-site cohort: Internal replication and extension. Schizophr Res. 2018;198:6-15. doi: 10.1016/j.schres.2017.05.013. PubMed PMID: 28549722; PubMed Central PMCID: PMCPMC5700873.

12. Roussos P, Giakoumaki SG, Zouraraki C, Fullard JF, Karagiorga VE, Tsapakis EM, et al. The Relationship of Common Risk Variants and Polygenic Risk for Schizophrenia to Sensorimotor Gating. Biol Psychiatry. 2016;79(12):988-96. doi: 10.1016/j.biopsych.2015.06.019. PubMed PMID: 26212897.

13. Greenwood TA, Lazzeroni LC, Maihofer AX, Swerdlow NR, Calkins ME, Freedman R, et al. Genome-wide Association of Endophenotypes for Schizophrenia From the Consortium on the Genetics of Schizophrenia (COGS) Study. JAMA Psychiatry. 2019;76(12):1274-84. doi: 10.1001/jamapsychiatry.2019.2850. PubMed PMID: 31596458; PubMed Central PMCID: PMCPMC6802253.

14. Watanabe A, Toyota T, Owada Y, Hayashi T, Iwayama Y, Matsumata M, et al. Fabp7 maps to a quantitative trait locus for a schizophrenia endophenotype. PLoS Biol. 2007;5(11):e297. doi: 10.1371/journal.pbio.0050297. PubMed PMID: 18001149; PubMed Central PMCID: PMCPMC2071943.

15. Shimamoto C, Ohnishi T, Maekawa M, Watanabe A, Ohba H, Arai R, et al. Functional characterization of FABP3, 5 and 7 gene variants identified in schizophrenia and autism spectrum disorder and mouse behavioral studies. Hum Mol Genet. 2014;23(24):6495-511. doi: 10.1093/hmg/ddu369. PubMed PMID: 25027319; PubMed Central PMCID: PMCPMC4240203.

16. Shimamoto-Mitsuyama C, Ohnishi T, Balan S, Ohba H, Watanabe A, Maekawa M, et al. Evaluation of the role of fatty acid-binding protein 7 in controlling schizophrenia-relevant phenotypes using newly established knockout mice. Schizophr Res. 2020;217:52-9. doi: 10.1016/j.schres.2019.02.002. PubMed PMID: 30765249.

17. Hirano S, Takeichi M. Cadherins in brain morphogenesis and wiring. Physiol Rev. 2012;92(2):597-634. doi: 10.1152/physrev.00014.2011. PubMed PMID: 22535893.

18. Sotomayor M, Gaudet R, Corey DP. Sorting out a promiscuous superfamily: towards cadherin connectomics. Trends Cell Biol. 2014;24(9):524-36. Epub 04/30. doi: 10.1016/j.tcb.2014.03.007. PubMed PMID: 24794279; PubMed Central PMCID: PMCPMC4294768.

19. Jaiganesh A, Narui Y, Araya-Secchi R, Sotomayor M. Beyond Cell-Cell Adhesion: Sensational Cadherins for Hearing and Balance. Cold Spring Harb Perspect Biol. 2018;10(9). doi: 10.1101/cshperspect.a029280. PubMed PMID: 28847902; PubMed Central PMCID: PMCPMC6008173.

20. Glover G, Mueller KP, Sollner C, Neuhauss SC, Nicolson T. The Usher gene cadherin 23 is expressed in the zebrafish brain and a subset of retinal amacrine cells. Mol Vis. 2012;18:2309-22. Epub 09/05. PubMed PMID: 22977299; PubMed Central PMCID: PMCPMC3441156.

21. Libe-Philippot B, Michel V, Boutet de Monvel J, Le Gal S, Dupont T, Avan P, et al. Auditory cortex interneuron development requires cadherins operating hair-cell mechanoelectrical transduction. Proc Natl Acad Sci U S A. 2017;114(30):7765-74. doi: 10.1073/pnas.1703408114. PubMed PMID: 28705869; PubMed Central PMCID: PMCPMC5544301.

22. Di Palma F, Holme RH, Bryda EC, Belyantseva IA, Pellegrino R, Kachar B, et al. Mutations in Cdh23, encoding a new type of cadherin, cause stereocilia disorganization in waltzer, the mouse model for Usher syndrome type 1D. Nat Genet. 2001;27(1):103-7. doi: 10.1038/83660. PubMed PMID: 11138008.

23. Watson C, Lind CR, Thomas MG. The anatomy of the caudal zona incerta in rodents and primates. J Anat. 2014;224(2):95-107. doi: 10.1111/joa.12132. PubMed PMID: 24138151; PubMed Central PMCID: PMCPMC3969054.

24. Swerdlow NR, Light GA. Sensorimotor gating deficits in schizophrenia: Advancing our understanding of the phenotype, its neural circuitry and genetic substrates. Schizophr Res. 2018;198:1-5. doi: 10.1016/j.schres.2018.02.042. PubMed PMID: 29525460; PubMed Central PMCID: PMCPMC6103885.

25. Toyoshima M, Akamatsu W, Okada Y, Ohnishi T, Balan S, Hisano Y, et al. Analysis of induced pluripotent stem cells carrying 22q11.2 deletion. Transl Psychiatry. 2016;6(11):e934. doi: 10.1038/tp.2016.206. PubMed PMID: 27801899; PubMed Central PMCID: PMCPMC5314118.

26. Toyoshima M, Jiang X, Ogawa T, Ohnishi T, Yoshihara S, Balan S, et al. Enhanced carbonyl stress induces irreversible multimerization of CRMP2 in schizophrenia pathogenesis. Life Sci Alliance. 2019;2(5):e201900478. doi: 10.26508/lsa.201900478. PubMed PMID: 31591136; PubMed Central PMCID: PMCPMC6781483.

27. Noben-Trauth K, Zheng QY, Johnson KR. Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. Nat Genet. 2003;35(1):21-3. Epub 2003/08/12. doi: 10.1038/ng1226. PubMed PMID: 12910270; PubMed Central PMCID: PMCPMC2864026.

28. Johnson KR, Tian C, Gagnon LH, Jiang H, Ding D, Salvi R. Effects of Cdh23 single nucleotide substitutions on age-related hearing loss in C57BL/6 and 129S1/Sv mice and comparisons with congenic strains. Sci Rep. 2017;7:44450. doi: 10.1038/srep44450. PubMed PMID: 28287619; PubMed Central PMCID: PMCPMC5347380.

29. Yasuda SP, Seki Y, Suzuki S, Ohshiba Y, Hou X, Matsuoka K, et al. c.753A>G genome editing of a Cdh23(ahl) allele delays age-related hearing loss and degeneration of cochlear hair cells in C57BL/6J mice. Hear Res. 2020;389:107926. doi: 10.1016/j.heares.2020.107926. PubMed PMID: 32101784.

30. Mianne J, Chessum L, Kumar S, Aguilar C, Codner G, Hutchison M, et al. Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. Genome Med. 2016;8(1):16. doi: 10.1186/s13073-016-0273-4. PubMed PMID: 26876963; PubMed Central PMCID: PMCPMC4753642.

31. Miyasaka Y, Shitara H, Suzuki S, Yoshimoto S, Seki Y, Ohshiba Y, et al. Heterozygous mutation of Ush1g/Sans in mice causes early-onset progressive hearing loss, which is recovered by reconstituting the strain-specific mutation in Cdh23. Hum Mol Genet. 2016;25(10):2045-59. doi: 10.1093/hmg/ddw078. PubMed PMID: 26936824.

32. Heermann T, Garrett L, Wurst W, Fuchs H, Gailus-Durner V, Hrabe de Angelis M, et al. Crybb2 Mutations Consistently Affect Schizophrenia Endophenotypes in Mice. Mol Neurobiol. 2019;56(6):4215-30. doi: 10.1007/s12035-018-1365-5. PubMed PMID: 30291584.

33. Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, et al. Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. Cell. 2018;174(4):1015-30 e16. doi: 10.1016/j.cell.2018.07.028. PubMed PMID: 30096299; PubMed Central PMCID: PMCPMC6447408.

34. Wallen-Mackenzie A, Dumas S, Papathanou M, Martis Thiele MM, Vlcek B, Konig N, et al. Spatio-molecular domains identified in the mouse subthalamic nucleus and

neighboring glutamatergic and GABAergic brain structures. Commun Biol. 2020;3(1):338. doi: 10.1038/s42003-020-1028-8. PubMed PMID: 32620779; PubMed Central PMCID: PMCPMC7334224.

35. Waldeck T, Wyszynski B, Medalia A. The relationship between Usher's syndrome and psychosis with Capgras syndrome. Psychiatry. 2001;64(3):248-55. doi: 10.1521/psyc.64.3.248.18467. PubMed PMID: 11708050.

36. McDonald C, Kenna P, Larkin T. Retinitis pigmentosa and schizophrenia. European psychiatry : the journal of the Association of European Psychiatrists. 1998;13(8):423-6. Epub 1998/12/01. doi: 10.1016/S0924-9338(99)80691-9. PubMed PMID: 19698660.

37. Domanico D, Fragiotta S, Cutini A, Grenga PL, Vingolo EM. Psychosis, Mood and Behavioral Disorders in Usher Syndrome: Review of the Literature. Med Hypothesis Discov Innov Ophthalmol. 2015;4(2):50-5. PubMed PMID: 26060830; PubMed Central PMCID: PMCPMC4458326.

38. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43. doi: 10.1038/s41586-020-2308-7.

39. Watanabe Y, Isshiki M, Ohashi J. Prefecture-level population structure of the Japanese based on SNP genotypes of 11,069 individuals. Journal of Human Genetics. 2020. doi: 10.1038/s10038-020-00847-0.

40. Collins SC, Mikhaleva A, Vrcelj K, Vancollie VE, Wagner C, Demeure N, et al. Large-scale neuroanatomical study uncovers 198 gene associations in mouse brain morphogenesis. Nat Commun. 2019;10(1):3465. doi: 10.1038/s41467-019-11431-2. PubMed PMID: 31371714; PubMed Central PMCID: PMCPMC6671969.

41. Van den Buuse M, Garner B, Koch M. Neurodevelopmental animal models of schizophrenia: effects on prepulse inhibition. Current molecular medicine. 2003;3(5):459-71. Epub 2003/08/29. doi: 10.2174/1566524033479627. PubMed PMID: 12942999.

42. Mitrofanis J. Some certainty for the "zone of uncertainty"? Exploring the function of the zona incerta. Neuroscience. 2005;130(1):1-15. Epub 2004/11/25. doi: 10.1016/j.neuroscience.2004.08.017. PubMed PMID: 15561420.

43. Fendt M, Li L, Yeomans JS. Brain stem circuits mediating prepulse inhibition of the startle reflex. Psychopharmacology (Berl). 2001;156(2-3):216-24. doi: 10.1007/s002130100794. PubMed PMID: 11549224.

44. Kita T, Osten P, Kita H. Rat subthalamic nucleus and zona incerta share extensively overlapped representations of cortical functional territories. J Comp Neurol. 2014;522(18):4043-56. doi: 10.1002/cne.23655. PubMed PMID: 25048050; PubMed Central PMCID: PMCPMC4198502.

45. Lindemann C, Krauss JK, Schwabe K. Deep brain stimulation of the subthalamic nucleus in the 6-hydroxydopamine rat model of Parkinson's disease: effects on sensorimotor gating. Behav Brain Res. 2012;230(1):243-50. doi: 10.1016/j.bbr.2012.02.009. PubMed PMID: 22330948.

46. Umemura A, Oka Y, Okita K, Matsukawa N, Yamada K. Subthalamic nucleus stimulation for Parkinson disease with severe medication-induced hallucinations or delusions. Journal of neurosurgery. 2011;114(6):1701-5. Epub 2011/03/08. doi: 10.3171/2011.2.JNS101261. PubMed PMID: 21375379.

47. Hawi Z, Tong J, Dark C, Yates H, Johnson B, Bellgrove MA. The role of cadherin genes in five major psychiatric disorders: A literature update. Am J Med Genet B Neuropsychiatr Genet. 2018;177(2):168-80. doi: 10.1002/ajmg.b.32592. PubMed PMID: 28921840.

48. Clifton EAD, Perry JRB, Imamura F, Lotta LA, Brage S, Forouhi NG, et al. Genome-wide association study for risk taking propensity indicates shared pathways with body mass index. Commun Biol. 2018;1(1):36. doi: 10.1038/s42003-018-0042-6. PubMed PMID: 30271922; PubMed Central PMCID: PMCPMC6123697.

49. Harr B, Karakoc E, Neme R, Teschke M, Pfeifle C, Pezer Z, et al. Genomic resources for wild populations of the house mouse, Mus musculus and its close relative Mus spretus. Sci Data. 2016;3(1):160075. doi: 10.1038/sdata.2016.75. PubMed PMID: 27622383; PubMed Central PMCID: PMCPMC5020872.

50. Tam WY, Cheung KK. Phenotypic characteristics of commonly used inbred mouse strains. J Mol Med (Berl). 2020. doi: 10.1007/s00109-020-01953-4. PubMed PMID: 32712726.

51. Mekada K, Hirose M, Murakami A, Yoshiki A. Development of SNP markers for C57BL/6N-derived mouse inbred strains. Exp Anim. 2015;64(1):91-100. doi: 10.1538/expanim.14-0061. PubMed PMID: 25341966; PubMed Central PMCID: PMCPMC4329520.

52. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science. 1997;276(5321):2045-7. doi: 10.1126/science.276.5321.2045. PubMed PMID: 9197268.

53. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, et al. Japanese Population Structure, Based on SNP Genotypes from 7003 Individuals Compared to Other Ethnic Groups: Effects on Population-Based Association Studies. The American Journal of Human Genetics. 2008;83(4):445-56. doi: https://doi.org/10.1016/j.ajhg.2008.08.019.

54. Dean B, Pavey G, Chai SW, Mendelsohn FAO. The Localisation and Quantification of Molecular Changes in the Human Brain Using in Situ Radioligand Binding and Autoradiography. Sydney: Gordon & Breach Science Publishers; 1999.

55. Kingsbury AE, Foster OJF, Nisbet AP, Cairns N, Bray L, Eve DJ, et al. Tissue pH as an indicator of mRNA preservation in human post-mortem brain. Molecular Brain Research. 1995;28(2):311-8. doi: 10.1016/0169-328x(94)00219-5.

56. Scarr E, Um JY, Cowie TF, Dean B. Cholinergic muscarinic M4 receptor gene polymorphisms: a potential risk factor and pharmacogenomic marker for schizophrenia. Schizophr Res. 2013;146(1-3):279-84. doi: 10.1016/j.schres.2013.01.023. PubMed PMID: 23490763.

57. Hill C, Keks N, Roberts S, Opeskin K, Dean B, MacKinnon A, et al. Problem of diagnosis in postmortem brain studies of schizophrenia. The American journal of psychiatry. 1996;153(4):533-7. Epub 1996/04/01. doi: 10.1176/ajp.153.4.533. PubMed PMID: 8599402.

58. Roberts SB, Hill CA, Dean B, Keks NA, Opeskin K, Copolov DL. Confirmation of the diagnosis of schizophrenia after death using DSM-IV: a Victorian experience. The Australian and New Zealand journal of psychiatry. 1998;32(1):73-6. Epub 1998/05/02. doi: 10.3109/00048679809062709. PubMed PMID: 9565186.

59. Zeng ZB. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci U S A. 1993;90(23):10972-6. doi: 10.1073/pnas.90.23.10972. PubMed PMID: 8248199; PubMed Central PMCID: PMCPMC47903.

60. Hayashi T, Iwata H. A Bayesian method and its variational approximation for prediction of genomic breeding values in multiple traits. BMC Bioinformatics. 2013;14(1):34. doi: 10.1186/1471-2105-14-34. PubMed PMID: 23363272; PubMed Central PMCID: PMCPMC3574034.

61. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157(4):1819-29. PubMed PMID: WOS:000168223400036.

62. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153(4):910-8. doi: 10.1016/j.cell.2013.04.025. PubMed PMID: 23643243; PubMed Central PMCID: PMCPMC3969854.

63. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature Protocols. 2013;8(11):2281-308. doi: 10.1038/nprot.2013.143. PubMed PMID: WOS:000326164100014.

64. Ohnishi T, Balan S, Toyoshima M, Maekawa M, Ohba H, Watanabe A, et al. Investigation of betaine as a novel psychotherapeutic for schizophrenia. Ebiomedicine. 2019;45:432-46. doi: 10.1016/j.ebiom.2019.05.062. PubMed PMID: WOS:000475860000044.

65. Zhang Q, Goto H, Akiyoshi-Nishimura S, Prosselkov P, Sano C, Matsukawa H, et al. Diversification of behavior and postsynaptic properties by netrin-G presynaptic adhesion family proteins. Molecular Brain. 2016;9(1):6. doi: 10.1186/s13041-016-0187-5.

66. Miyasaka Y, Suzuki S, Ohshiba Y, Watanabe K, Sagara Y, Yasuda SP, et al. Compound heterozygosity of the functionally null Cdh23v-ngt and hypomorphic Cdh23ahl alleles leads to early-onset progressive hearing loss in mice. Experimental animals. 2013;62(4):333-46.

67. Mashiko H, Yoshida AC, Kikuchi SS, Niimi K, Takahashi E, Aruga J, et al. Comparative anatomy of marmoset and mouse cortex from genomic expression. J Neurosci. 2012;32(15):5039-53. doi: 10.1523/JNEUROSCI.4788-11.2012. PubMed PMID: 22496550; PubMed Central PMCID: PMCPMC6622108.

68. Shimogori T, Abe A, Go Y, Hashikawa T, Kishi N, Kikuchi SS, et al. Digital gene atlas of neonate common marmoset brain. Neurosci Res. 2018;128:1-13. doi: 10.1016/j.neures.2017.10.009. PubMed PMID: 29111135.

69. Boyle EA, O'Roak BJ, Martin BK, Kumar A, Shendure J. MIPgen: optimized modeling and design of molecular inversion probes for targeted resequencing. Bioinformatics. 2014;30(18):2670-2. Epub 2014/05/29. doi: 10.1093/bioinformatics/btu353. PubMed PMID: 24867941; PubMed Central PMCID: PMCPMC4155255.

70. 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):246-50. doi: 10.1038/nature10989. PubMed PMID: 22495309; PubMed Central PMCID: PMCPMC3350576.

Figure legends

Figure 1: QTL mapping for PPI and expression analysis of *Cdh23*. (**A**) Composite interval mapping revealed high LOD scores on chromosome 10 for the 86 dB PPI, (**B**) which peaked at a synonymous coding and splice-site variant c.753 G>A (rs257098870) in the *Cdh23* gene. (**C**) Bayesian multiple QTL mapping for the 86 dB PPI consistently showed the same variant with a posterior probability of 1, underscoring the QTL effect. (**D**) RNA *in situ* hybridization showed that *Cdh23* is primarily expressed in subthalamic (including the zona incerta and subthalamic nucleus) and pontine regions in both B6N and C3HN mice. Representative image from 3-week-old B6N and C3HN (scale bar = 500 µm). (**E**) *CDH23* transcript expression was shown to be developmentally regulated across different brain regions in humans (http://development.psychencode.org/). (**F**) *In vitro* neuronal differentiation from hiPSCs. Expression analysis was performed in 4 hiPSC lines, neurospheres derived from each hiPSC line in triplicate, and neurons derived from one of the hiPSC lines in triplicate.

Figure 2: *Cdh23* c.753G allele knock-in in the B6N genetic background dampened the PPI, rescued skipping of exon 9 and showed intact hearing acuity. (**A**) Alternative splicing resulting from *Cdh23* c.753G>A in the pontine and subthalamic regions was compared between *Cdh23* c.753G allele knock-in, *Cdh23* c.753A allele littermates (both litters of heterozygous intercross), and inbred B6N and C3HN mice using reverse RT-PCR. (**B**) Prepulse inhibition levels at different prepulse levels were tested in *Cdh23* c.753G allele knock-in mice in the B6N genetic background in comparison to *Cdh23* c.753A allele littermates (13 weeks old). The results showed reduced PPI levels in G allele knock-in mice when compared to the A allele littermates in all the prepulse levels tested (age = 13-14 weeks, AA genotype; n = 33, GG genotype; n = 34). Statistical significance was determined

using the Holm-Sidak method, with alpha = 0.05. (**C**) The ABR thresholds at 8, 16, and 24 kHz were determined to test the hearing acuities in *Cdh23* c.753G allele knock-in mice in comparison to *Cdh23* c.753A allele litter mates (age = 13 weeks, AA genotype; n = 18, GG genotype; n = 16). No differences in ABR thresholds were observed in either genotype, indicating intact hearing. (**D**) ABR measurements for the same frequencies in 1-year-old mice showed a lower ABR threshold for the G allele knock-in mouse and a higher threshold for A allele carriers, indicating hearing loss (AA genotype; n = 6, GG genotype; n = 9). Statistical significance was determined using an unpaired two-tailed Student's *t*-test (**E**). Scanning electron microscope (SEM) images of cochlear hair cells showed intact stereociliary architecture in both genotypes at 8 weeks, and the stereociliary morphology was preserved in 6-month-old *Cdh23* c.753G allele knock-in mice in the B6N genetic background when compared to the *Cdh23* c.753A allele littermates. The data are represented as the mean ± SEM. *P < 0.05, **P < 0.01, ***P < 0.001

Figure 3: *Cdh23* transcript expression and genetic variants in *CDH23*. (**A**) Digital PCR-based *Cdh23* transcript level estimations in pontine and subthalamic regions from *Cdh23* c.753 G allele knock-in mouse and *Cdh23* c.753 A allele littermates revealed significantly lower expression in G allele carriers in both brain regions (n = 6 per group; age 4-6 weeks). Statistical significance was determined using the Holm-Sidak method, with alpha = 0.05. (**B**) Although, in inbred B6N and C3HN mice, no differences in *Cdh23* transcript levels were observed for either genotype in adults (n = 6 per group; age 4-6 weeks), (**C**) significant differences in the expression pattern of *Cdh23* transcript across the development (E16.5, embryonic day 16.5; P0, at birth; P7, postnatal day 7) was observed between the genotypes. Whole cortex for E16.5 (B6N; n = 14, C3HN; n = 21), P0 (B6N; n = 16, C3HN; n = 16), and P7 (B6N; n = 16, C3HN; n = 17) were used for analysis. Statistical significance was determined using an unpaired two-tailed Student's *t*-test (**D**) Novel rare variants in *CDH23*

identified exclusively in schizophrenia cases are shown on the protein domains (see supplementary figure 11A). Indicating the level of consensus from the annotation tools, the red color denotes variants predicted to be deleterious by ≥ 4 tools, pink color denotes variants predicted to be deleterious by ≤ 3 tools, and the blue color indicates the variants not predicted to be deleterious. Additionally, the green color (**D, E**) indicates the syntenous variant for *Cdh23* c.753G>A, which was observed in humans (rs769896655; c.753G>A/T; P251P), and another variant that affects the same amino acid. (**F**) The variant major allele G for rs769896655 was shown to be conserved across the species. (**G**) *CDH23* transcript expression analysis showed no difference between postmortem brain samples from Brodmann area 8 (BA8) of schizophrenia (n = 50) and control (n = 71). The median expression values are shown. Statistical significance was determined using the Mann–Whitney U-test (two-tailed). (**H**) Allele-specific expression of the variant rs769896655 in the *CDH23* transcript was tested in hair follicles (n = 1 in triplicate), peripheral blood samples (n = 1 in triplicate), hiPSCs, hiPSC-derived neurospheres, and neurons (early neurons; day 7, mature neuron; day 30 of differentiation) from a healthy subject heterozygous for the variant. For hiPSCs, four lines were derived from one sample and were tested in duplicate. A preferential inclusion of the G allele in the *CDH23* transcript was observed only from peripheral blood samples and hiPSCs. Statistical significance was determined using an unpaired two-tailed Student's *t*-test. The data are represented as the mean ± SEM; *P < 0.05, **P < 0.01, ***P < 0.001.

Table1: Association analysis of *CDH23* variant rs769896655 with schizophrenia

Variant	Subjects	Sample size (n)	Alternate allele (A)	Reference allele (G)	*P*-value	Odds ratio	95% Confidence interval	Minor allele frequency
***CDH23* rs769896655 (Pro251Pro) chr10:71570918 G/A (GRCh38)**	**Schizophrenia**	**2,011**	**4**	**4,018**				**0.00099**
	Controls	**11,759**	**6**	**23,512**				**0.00026**
	Controls (RIKEN) (Japanese)	2,170	2	4,338	**0.0458**	3.901	1.241 - 14.28	0.00046
	ToMMO 8.3KJPN (Japanese)	8,380	3	16,757				0.00018
	HGVD (Japanese)	1,209	1	2,417				0.00041

RIKEN; Schizophrenia and controls sampled from Honshu area of Japan (the main island of Japan)
ToMMO 8.3KJPN; Tohoku Medical Megabank Organization (https://jmorp.megabank.tohoku.ac.jp/202008/downloads#variant)
HGVD; The Human Genetic Variation Database Kyoto (http://www.hgvd.genome.med.kyoto-u.ac.jp/index.html)

List of supplementary tables

Supplementary table 1: Strain-specific single nucleotide variants located on the previously reported loci for PPI, selected for genotyping

Supplementary Table 2: QTL mapping performed for 86 dB PPI by composite interval mapping with LOD score and percentage of variance explained for the phenotype (R2) in 1,010 F2 mice derived from F1 parents (B6NjC3Hj), both combined and dichotomized for gender

Supplementary Table 3: Bayesian multiple QTL mapping performed for 86 dB-PPI with posterior probabilities across additive and dominant models along with percentage of variance explained (R2) in 1,010 F2 mice derived from F1 parents (B6NjC3Hj), both combined and dichotomized for gender

Supplementary Table 4: Single nucleotide variants, which were distinct between the C57BL/6NJ and C3H/HeJ mouse strains, within the chromosome 10 QTL region from the sequence variation database of mouse strains (https://www.sanger.ac.uk/science/data/mouse-genomes-project/; REL-1303- GRCm38)

Supplementary Table 5: Variant allele frequency of the *CDH23* rs769896655 among different populations

Supplementary Table 6: Association analysis of *CDH23* variant rs769896655 with schizophrenia

Supplementary Table 7: Sequence for short guide RNA (sgRNA) and single-stranded-oligodeoxynucleotide (ssODN) for generating CRISPR/Cas9n mediated knock-in mouse

Supplementary Table 8: Primer sequences for the gene expression analysis

Supplementary Table 9: Molecular inversion probes sequences targeting the coding exons and the flanking exon-intron boundaries of *CDH23*

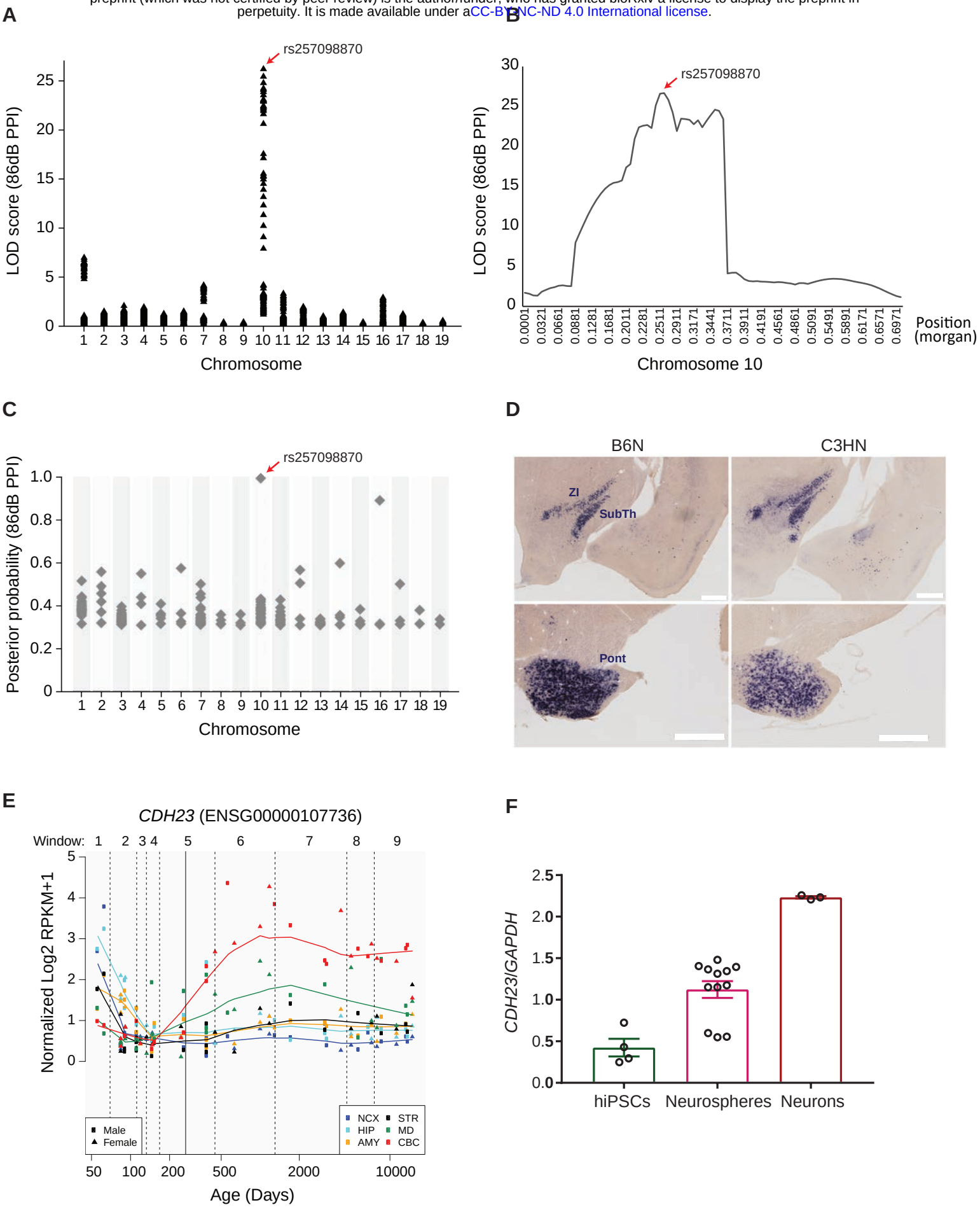

Figure 1

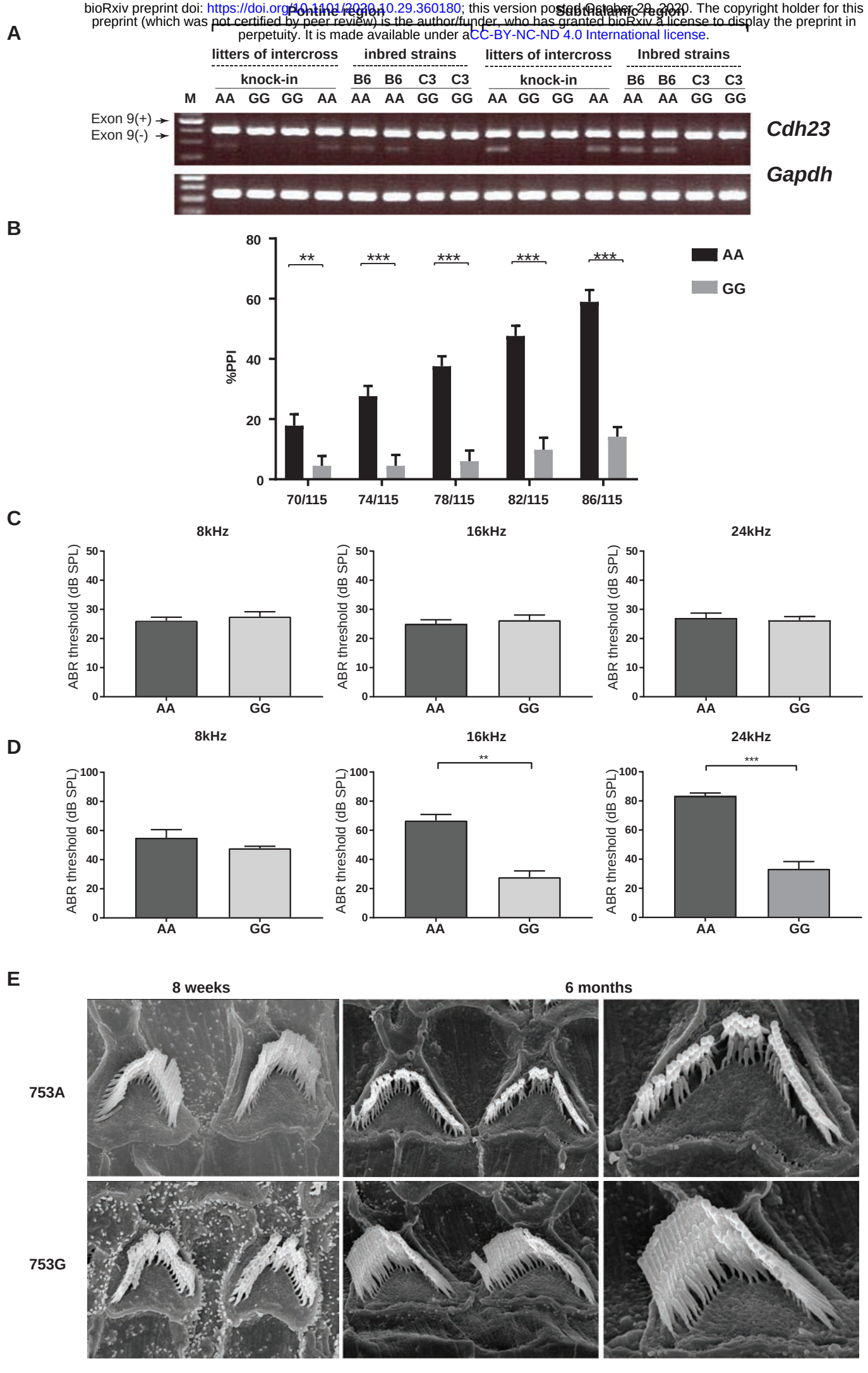

Figure 2

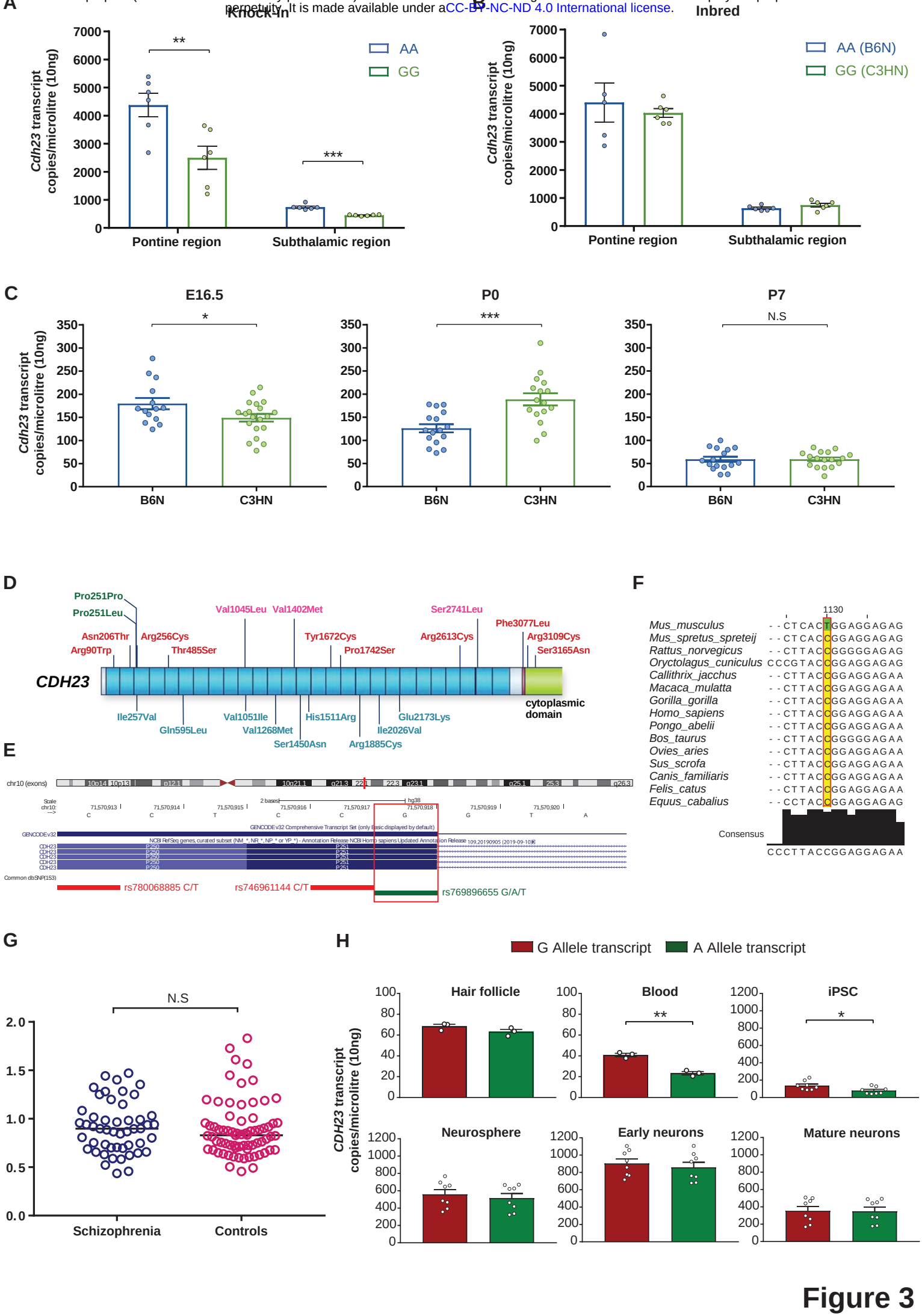

Figure 3