Loss of δ–catenin function in severe autism

Tychele N. Turner1,2,3, Kamal Sharma4, Edwin C. Oh5, Yangfan P. Liu5, Ryan L. Collins6, Maria X. Sosa1,3, Dallas R. Auer1,3, Harrison Brand6,7, Stephan J. Sanders3,8, Daniel Moreno-De-Luca1,9, Vasyi Pluhir1,3, Teri Plona10, Kristen Pike10, Daniel R. Soppet10, Michael W. Smith11, Sau Wai Cheung12, Christa Lese Martin1,13, Matthew W. State3,8, Michael E. Tolkowski6,7, Edwin Cook3,4, Richard Huganir4, Nicholas Katsanis5 & Aravinda Chakravarti1,3

Autism is a multifactorial neurodevelopmental disorder affecting more males than females; consequently, under a multifactorial genetic hypothesis, females are affected only when they cross a higher biological threshold. We hypothesize that deleterious variants at conserved residues are enriched in severely affected patients arising from female–enriched multiplex families with severe disease, enhancing the detection of key autism genes in modest numbers of cases. Here we show the use of this strategy by identifying missense and dosage sequence variants in the gene encoding the adhesive junction–associated δ–catenin protein (CTNND2) in female–enriched multiplex families and demonstrating their loss–of–function effect by functional analyses in zebrafish embryos and cultured hippocampal neurons from wild–type and Ctnd2 null mouse embryos. Finally, through gene expression and network analyses, we highlight a critical role for CTNND2 in neuronal development and an intimate connection to chromatin biology. Our data contribute to the understanding of the genetic architecture of autism and suggest that genetic analyses of phenotypic extremes, such as female–enriched multiplex families, are of innate value in multifactorial disorders.

Autism is a common neurodevelopmental disorder with a profound sex-bias: four times more males than females are affected whereas disease recurrence risk to siblings of autistic females is larger than to siblings of affected males. Both features can be explained through autism’s multifactorial inheritance where females are affected at higher biological thresholds of an underlying liability than males. Under this model, females escape the effect of deleterious mutations unless the alleles are severe and at key developmental steps. To accelerate discovery, we examine families with highest recurrence risk and, consequently, probably enriched for severe mutations in such genes. We hypothesize that one group of families that have this property, and yet are underrepresented in autism sequencing efforts, are those with two or more severely affected females (female–enriched multiplex families (FEMFs)). The first genes discovered in autism were through syndromes (Supplementary Table 1), such as Rett and fragile X syndromes. Today, genomic analyses have definitively identified 12 genes, from an estimated 500 (ref. 4), with an excess of de novo or segregating mutations in typical isolated cases that are overwhelmingly male (Supplementary Table 1). Given such heterogeneity, it may be crucial to identify those genes whose mutations impart the greatest autism risk. Increased recurrence risk is associated with lower incidence (the ‘Carter’ effect), since any rare class must arise from higher genetic liability (Fig. 1a). Consequently, gene discovery in epidemiologically rarer classes, namely female gender, high phenotypic severity and familial cases, may be fruitful; this is further enhanced if we increase the genetic load by considering individuals who have all three features.

These genetically loaded cases have either a greater number or frequency of deleterious alleles that are probably severe coding variants. This prediction arises from our studies of Hirschsprung disease, a neurodevelopment disorder of enteric nervous system ganglioneurosis. Hirschsprung disease is a multifactorial disorder with a sex ratio of 4:1 in favour of males and whose risk factors are gender, phenotypic severity, and familiality. Although more than 15 genes for Hirschsprung disease have been identified, the major gene encodes the receptor tyrosine kinase activity of β-catenin protein (β-catenin function in severe autism).

Figure 1 | Genetic features of a sex-dependent multifactorial model.

a. Hypothetical sex-dependent liability distributions for autism under a multifactorial model of inheritance with a fixed biological threshold for affection. b. Percentage of patients having Hirschsprung disease with damaging coding mutations within different risk classes characterized by gender, segment length and familiarity. The risk class is labelled 3, 2, 1, 0 and is an additive score based on the number of factors with higher risk (female, long segment, multiplex); it comprises 13, 46, 60 and 55 patients, respectively (proportion trend test, $P = 3.1 \times 10^{-6}$).
kinase RET, which harbours numerous rare loss-of-function coding and one common enhancer variant. We estimated the proportion of 174 patients with Hirschsprung disease with damaging RET coding variants conditional on their having 3, 2, 1 or 0 risk factors, where higher risk categories were female gender, long segment aganglionosis and family history (Fig. 1b), to show that rare classes are significantly associated with a higher proportion of deleterious alleles, varying linearly between 46% and 2% from the highest to lowest risk class (P = 3.1 × 10−6); the non-coding variant had the reverse trend. Therefore, exome sequencing in autism can be similarly efficient in FEMFs. Since female incidence of autism is 0.0016, fewer than 10% of families are multiplex and fewer than 10% are severe, FEMFs have a crude incidence of less than 1.6 × 10−5 and represent a rare autism disorder enriched for deleterious coding variants.

Here we demonstrate the utility of this strategy by exome sequence analyses of 13 unrelated females and identifying 18 candidate genes of which at least four, CYFIP1, DLG1, PLXNA3 and CTNND2, are of interest to autism aetiology. We have evaluated one of them, CTNND2 (the δ-2-catenin gene encoding the δ-catenin protein), in depth using a combination of genetic, genomic and functional studies to show that (1) CTNND2 harbours a significant excess of deleterious missense and copy number variants (CNVs) in autism; (2) these variants, by functional testing, are loss-of-function and affect Wnt signalling; (3) expression of CTNND2 is highest in the fetal brain and is highly correlated with other autism genes; and (4) CTNND2 correlated genes are enriched for chromatin and histone modification, as well as dendritic morphogenesis, functions. These results are consistent with the roles of CTNND2 in the formation of dendritic spines, and the regulation of β-(δ)-catenin in neurons. Given the recent finding of de novo autism mutations in pathways regulating β-catenin (Supplementary Table 1), loss-of-function of CTNND2 is probably rate-limiting for dendritic morphogenesis and maintenance.

Exome sequencing of females with autism
We sampled 13 unrelated females, negative for deleterious variants in MECP2, from multiplex families who had severe autism (measured using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule). Proband exomes were sequenced and analysed with sequence data from 71 European females (1000 Genomes Project; Extended Data Fig. 1). To identify pathogenic alleles, we focused on missense variants absent in public databases (dbSNP129, 1000 Genomes Project) and conserved to zebrafish, nonsense and canonical splice site variants. This led to 3,090 variants of interest in the combined 84 exomes within 2,516 genes, with 447 of these having two or more variants of interest; among them, the 13 autism cases harboured two or more variants of interest in 24 genes of which 18 reached significance (P < 1 × 10−5) (Supplementary Table 2). By searching their expression profiles (Supplementary Table 3), we identified four genes, with an excess of deleterious alleles, as candidates: CYFIP1, DLG1, PLXNA3 and CTNND2. On the basis of our previous genome-wide association study implicating chromosome 5p15, we followed up CTNND2 at this locus.

**CTNND2 as a novel autism gene**
CTNND2 harboured two deleterious variants, G34S and R713C, both of which were absent in 3,889 European controls (1000 Genomes Project and Exome Variant Server); G34S was present at a frequency of 5.3 × 10−8 in 1,869 African ancestry samples (Exome Variant Server) and in one Luhyan sample (NA19020) (Extended Data Fig. 2). To estimate their frequency, we genotyped 10,782 samples from the HapMap and autism collections: the only additional individuals with G34S were an affected female and her mother (SSC02696, SSC03276) from the Simons Simplex Collection (SSC). Principal component analysis on polymorphism data from individuals with G34S found that our autism cases were not of African ancestry, identifying a new ancestral origin for G34S (Extended Data Fig. 3). For R713C, only our FEMF samples were heterozygous. Next-generation CTNND2 sequencing in 362 additional females with autism (Extended Data Fig. 4) identified a total of seven variants (G34S, R713C and five new variants: P189L, P224L, G275C, R454H, T862M), of which four (G34S, G275C, R713C, T862M) were conserved to zebrafish (Fig. 2a and Supplementary Table 4). We also identified Q507P in an autistic male from 170 SSC probands. An identical analysis of 379 European ancestry control samples (1000 Genomes Project) yielded

---

**Figure 2 | Missense variants in human δ-catenin and their effect on protein function in vivo. a, CTNND2 annotated with validated missense mutations in autism patients; G34S, G275C, Q507P, R713C, T862M variants are conserved to zebrafish. b, Expression of two CTNND2 zebrafish orthologs (ctnnd2a, ctnnd2b) in development. Abrerrant phenotypes are observed with ctnnd2b (the only orthologue expressed at these gastrulation time points) morpholino knockdown only at key stages of gastrulation (50% epiboly, 75% epiboly, bud). Elongation factor alpha (efsla11) is shown as a control for ubiquitous expression. c, Representative lateral and dorsal images of class I and class II ctnnd2b morphants (2 ng morpholino) at the eight- to ten-somite stage (A) is marked for each condition.**
three variants after validation (R330H, D465N, A482T), one conserved to zebrafish. On aggregate, variants at these conserved CTNND2 residues are significantly more frequent in autism than in controls ($P = 0.04$ versus 1000 Genomes Project; $P = 7.8 \times 10^{-4}$ versus Exome Variant Server).

We next assessed whether CNVs within CTNND2 were enriched in autism. First, from the literature, we identified six deletions and one duplication. Second, we identified two deletions and one duplication from the Emory University and Baylor College of Medicine clinical cytogenetics laboratories. Third, from the Autism Genetic Resource Exchange (AGRE), we identified two previously unreported valid deletions (Extended Data Fig. 5). Therefore, we detected 12 CNVs (ten deletions, two duplications), seven overlapping one or more exons (Fig. 3 and Supplementary Table 5). As a control, we searched the Database of Genomic Variants to identify 33 variants, with only two overlapping exons (58.3% in our 12 CNVs versus 6.1% in the Database of Genomic Variants; $P = 5 \times 10^{-4}$). This significant excess of exon-disruptive deletions suggests CTNND2 haploinsufficiency in autism. Most of our patients had an autism diagnosis; however, some probands were referred with a diagnosis of neurodevelopment disorder. To test whether CTNND2 CNVs may be enriched in neurodevelopment disorders generally, we assessed CNVs from 19,556 independent cases referred for clinical diagnostic studies and 13,098 controls from population-based studies.

Considering all dosage imbalances, we observed 25 instances in cases and 13 in controls, corresponding to an odds ratio of 5.9 ($P = 4.10 \times 10^{-4}$) (Extended Data Fig. 6 and Supplementary Table 6). The impact of loss-of-function (deletions, unbalanced translocations) mutations at this locus is significant, with an odds ratio of 14.7 ($P < 0.001$), implicating these as hypomorphic (Fig. 2d). One variant (R713C) was functionally null while G275C and T862M were benign, and all four controls were benign, demonstrating specificity. To preclude the possibility of mRNA toxicity, we injected mutant mRNA corresponding to all alleles and observed no significant differences in the gastrulation phenotypes (Fig. 2e).

To replicate these findings with an in vivo assay querying Wnt signalling earlier in development, we assessed the consequences of ctnnd2b suppression on chordin expression during epiboly, whose ectopic expression is known in Wnt mutants\(^\text{43}\). Consistent with the role of chordin in Wnt-dependent dorsalization\(^\text{14}\), we observed shortening and widening of the chordin expression domain as well as loss of anterior-specific expression fields in ctnnd2b morphants (Extended Data Fig. 7b); this phenotype could be rescued by wild-type mRNA. Further, testing of two control alleles that scored benign in our mid-somatic assays (A482T, G810R) showed significant rescue ($P < 0.001$); the hypomorphic allele G34S rescued chordin expression to a level significantly worse than wild-type mRNA rescue ($P < 0.001$), while the null allele R713C did not rescue chordin expression (Extended Data Fig. 7c). Since CTNND2 can bind CTNNB1 (ref. 15), we tested this interaction with mutant CTNND2. Expression of green fluorescent protein (GFP)-tagged CTNND2 and Flag-tagged CTNNB1 revealed that wild-type CTNND2 could immunoprecipitate CTNNB1; however, its interaction with CTNNB1 was diminished upon expression of G34S or R713C (Extended Data Fig. 7d), suggesting in vivo Wnt phenotypes may result from attenuated CTNNB1–CTNND2 interaction.

Finally, we asked if these major CTNND2 sequence variants could affect neuronal circuitry by using a well-established in vitro model system. Dendritic spines are the primary sites for excitatory synapse formation, and their dysregulation underlies many neuropsychiatric disorders\(^\text{46}\). To test if CTNND2 variants interfere with development and maintenance of spines, we prepared primary hippocampal neurons from embryonic day (E)18 rat embryos and introduced either GFP or GFP fusion to wild-type CTNND2 or to its mutant variants at day in vitro (DIV)8. At DIV15, neurons were fixed and analysed to assess spine density. We found that wild-type CTNND2 had a significantly higher spine density than GFP controls\(^\text{27}\). However, neurons expressing G34S had a significantly lower spine density than those expressing GFP or wild-type CTNND2. Neurons expressing R713C, on the other hand, had the same spine density as those expressing GFP but significantly less than the one that expressed wild-type CTNND2, suggesting a loss-of-function effect. In contrast, the A482T polymorphism had an effect similar to wild-type CTNND2 (Extended Data Fig. 8). To test if observed changes in spine density reflected changes in excitatory synapse number in the networks, we analysed excitatory synapses: that is, overlapping region between postsynaptic marker PSD95 and presynaptic marker vGlut1 in mouse hippocampal neurons at DIV14 (Fig. 4A). As with spine density, we found an increase in excitatory synapse number in neurons that overexpressed wild-type but not mutant CTNND2. Further, loss-of-function of CTNND2 led to a decrease in overall excitatory synapse density, as well as active synapses that expressed the GluA subunit of the AMPA (z-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid)-type glutamate receptors (Fig. 4B, C). Taken together, these results

![Figure 3](image-url)
null mutants have a significant reduction in synapse density. Synapses are of which PAX6 is the most biologically significant. To understand expression of CTNND2 in the developing human brain suggest that CTNND2 is critical to the formation and/or maintenance of synapses, in accord with other studies. Moreover, unlike wild-type CTNND2, the tested mutants failed to rescue the reduction in synapse density in CTNND2 null background, demonstrating loss-of-function. Therefore, G34S and R713C impair development and/or the maintenance of mammalian neural circuitry.

Expression of CTNND2 in the developing human brain
To understand CTNND2 expression, we tested mRNA levels in 16 adult and eight fetal human tissues: CTNND2 expression was highest in the fetal brain (20× the adult brain) (Extended Data Fig. 9). Therefore, we used the Allen Brain Atlas of the Developing Human Brain microarray data to identify other CTNND2 co-expressed genes. We used the data normalized to 17,630 genes and linear regression on age and brain regions for estimating Pearsonian correlations between CTNND2 and all other genes (absolute correlation > 0.3, \( P = 2.84 \times 10^{-5} \) given multiple comparisons). First, we performed pathway analysis on the 826 positively and 662 negatively correlated genes (Supplementary Table 7). The positive set was significantly enriched for genes encoding proteins localized to the cytoskeleton, cell junction, neuronal projection, with GTPase regulatory activity, and functioning in cell morphogenesis, chromatin modification, neuronal development and neuron projection formation. Of these, the role of CTNND2 in dendritic development and spine morphogenesis is known as well as its involvement in actin dynamics and GTPase regulatory activity. However, its role in chromatin modification is novel. The closest known function of CTNND2 to chromatin is based on CTNND2 binding to ZBTB33 (ref. 22), a protein regulating transcription and Wnt pathway genes, and its possible nuclear localization and function. Second, we searched for transcription factors that may regulate CTNND2: among the correlated genes we identified 75 of which PAX6 is the most biologically significant. A PAX6 mutant rat show autism-related features and genetic variation disrupting PAX6 has been identified in individuals with autism. Also, Pax6 can regulate Ctnnd2 expression in cells, including the binding of Pax6 to its promoter.

We searched the correlated genes for autism (https://gene.sfari.org/autdb/) and intellectual disability candidates (Supplementary Table 7). Of 529 autism genes, 71 (61 positively, 10 negatively) were significantly correlated with CTNND2, representing significant enrichment (\( P = 2.83 \times 10^{-5} \)). Next, we examined the correlations between these 71 genes and CTNND2 (Fig. 5a) to find an intimate relationship between CTNND2 and autism genes. To interrogate the function of the 61 positively correlated genes, we again performed pathway analyses (Fig. 5b) to find significant enrichment of genes involved in dendrite morphogenesis (\( P = 2.96 \times 10^{-10} \)). PDLIM5, MAP2, SHANK1, CDKL5, DLG4 as well as chromatin modification (\( P = 2.96 \times 10^{-12} \)), HDAC3, HUWE1, CREBBP, EP300, YEATS2, EP400, ATXN7, HCF1C, ARID1B, NSD1).

Discussion
Our studies strongly implicate δ-2-catenin (CTNND2) as a critical gene in autism and an important neurodevelopmental protein given its role in FEMFs, functional association with other autism genes, cri-du-chat syndrome and other diseases. Clearly, CTNND2 haploinsufficiency is common in autism and strongly associated with neurodevelopment disorder generally. Nevertheless, in the general population, the frequency of disease alleles we discovered is low (3.9 \( \times \) 10\(^{-4} \) and 8.0 \( \times \) 10\(^{-4} \) in individuals of European and African ancestry, respectively, in Exome Variant Server), consistent with their deleterious functional effects.

CTNND2 is a plakoglobin/armadillo family member with identity to PKP4, CTNND1 and ARVCF. The armadillo domain is a key part of the protein that binds cadherins, β-catenin, presenilins 1 and 2 (ref. 32) and sphingosine kinase. It also harbours a coiled-coil domain, a polyproline tract at amino acids 219–224 where src receptor kinases bind, and a PDZ domain at the carboxy (C) terminus, which can bind Discs large homologue 4 (ref. 35) and erbin. These features suggest that CTNND2 is important in neuronal actin dynamics and the cytoskeleton, as also supported by observations of induced branching of dendrite-like processes and enhanced dendrite morphogenesis by CTNND2 overexpression. Importantly, CTNND2 can directly bind to actin and cortactin, and act on the rho family to induce filopodia

![Figure 4](https://example.com/f4.png)

**Figure 4 | δ-Catenin is critical for maintaining functional neuronal networks.** A. Gain of function. Overexpression of CTNND2 leads to an increase in the number of excitatory synapses. Primary dendrites from neurons transfected with GFP alone, GFP fusion with wild-type (WT) CTNND2, or mutant isoforms, and immunolabelled with vGluT1 and PSD95. B. Quantification of number of PSD95 + vGluT1-positive puncta per 100 μm of dendritic length (\( N = 12 \) each). C. Loss of function. Neurons from Ctnnd2 null mutants have a significant reduction in synapse density. Synapses are identified as puncta with PSD95 and vGluT1 overlap, KO, knockout. B. Quantification of the number of PSD95 + vGluT1-positive puncta per 100 μm of dendritic length (\( N = 13 \) each). C. Alternatively, neurons were immunolabelled with GluA and vGluT1 to identify active functional excitatory synapses. D. Quantification of the number of GluA + vGluT1-positive puncta per 100 μm of dendritic length (\( N = 15 \) each). E. Rescue of loss of function. WT CTNND2 but not its mutant isoforms can rescue the loss of phenotype in neurons from CTNND2 null mutants. Primary dendrites from neurons transfected with GFP alone, GFP fusion with wild-type CTNND2, or mutant isoforms, and immunolabelled with vGluT1 and PSD95. B. Quantification of the number of PSD95 + vGluT1-positive puncta per 100 μm of dendritic length (\( N = 14 \) each). Colour used for merged panels are GFP (green), PSD95 (red), GluA (Red) and vGluT1 (blue). Student’s t-tests were conducted: \( \ast \) \( P < 0.05 \) and \( \ast \ast \) \( P < 0.001 \), respectively. Error bars, s.e.m.
within neurons\(^{29}\) and increase the number of dendritic spines\(^{37}\). Finally, we demonstrate a role of CTNND2 in canonical Wnt signalling through zebrafish analyses: although the precise mechanism is not understood, it can bind to proteins (GSK-3\(\beta\), ZBTB33) that regulate Wnt signalling\(^{23}\) and transcription\(^{24}\), in concert with CTNNB1. The other novel CTNND2 function we implicate is its possible role in the nucleus, through its interaction with HDAC3 (Fig. 5). This is not unexpected, since CTNND2 can affect gene expression after nuclear translocation\(^{24}\). Furthermore, the armadillo family member p120ctn interacts with ZBTB33 (Kaiso) and the NCoR co-repressor complex containing HDAC3 (refs 22, 24, 38, 39). Thus, we hypothesize that CTNND2 may be a nucleo-cytoplasmic protein whose autism effect may arise from its cytoplasmic or nuclear loss-of-function or both.

Published Ctnnd2 knockout mice\(^{8,18,19}\), and our analyses of their dendritic spines, give clues to the role of CTNND2 in autism and in cognition. Homozygote mice exhibit structural and functional abnormalities at the synapse, as well as impaired spatial learning and fear conditioning\(^{18,19}\), with reduced levels of PSD-95, \(\beta\)-catenin associated with cadherin, and N-cadherin. The interaction of PSD-95 with CTNND2 was discovered as an important linkage to AMPA receptor binding protein and glutamate receptor interacting protein (GRIP)\(^{35}\). Our results confirm that CTNND2 is required for the maintenance of spine structures in vivo\(^{39}\), and stability of some key components of the synaptogenic machinery such as N-cadherins and PSD95 (refs 8, 18). We show that loss of spines and reduction in total levels of synaptic proteins in null mice reflect reduction in the number of functional excitatory synapses at the subcellular level. Interestingly, acute loss of \(\delta\)-catenin in vitro impairs activity-dependent formation of spines\(^{40}\), reinforcing its importance in formation and maintenance of synaptic structures and cognitive functions.

Studying FEMFs is unconventional for a complex disease where most mutations have small effects. Nevertheless, our data suggest that modest numbers of samples of rare extreme phenotypes, in contrast to large numbers of typical cases, can be important. Note, we identified 18 candidate genes among which at least three others are worthy of follow-up: CYFIP\(1\) is in a 15q11-13 autism duplication, has altered expression in autism patients, interacts with FMRP and is involved in regulating dendritic spines through translational inhibition and actin dynamics\(^{41}\); DLGI is a multi-scaffolding postsynaptic density protein lying within a 3q29 autism and/or intellectual disability deletion; PLXNA3 is known to alter dendritic spines and is a receptor for SEMA5A\(^{42}\), another autism gene\(^{40}\). The broader FEMFs hypothesis can thus be tested by sequencing larger numbers of cases for identifying genes critical to early brain development.

Figure 5 | Gene expression correlation between CTNND2 and known autism genes. a. Plot of all autism genes significantly (positive and negative) correlated with CTNND2 in the developing human brain (microarray data from http://www.brainspan.org). b. Pathway analysis of the autism genes positively correlated with \(\delta\)-catenin reveals significant enrichment of genes involved in chromatin modification and dendrite morphogenesis.
Online Content
Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 January 2014; accepted 5 January 2015.

Published online 25 March 2015.

1. Kogan, M. D. et al. Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. Pediatrics 124, 1395–1403 (2009).
2. Jorde, L. B. et al. Complex segregation analysis of autism. Am. J. Hum. Genet. 49, 932–938 (1991).
3. Abrahams, B. S. & Geschwind, D. H. Advances in autism genetics: on the threshold of a new neurobiology. Nature Rev. Genet. 9, 341–355 (2008).
4. Ronemus, M., Lesslov, I., Levy, D. & Wigler, M. The role of de novo mutations in the genetics of autism spectrum disorders. Nature Rev. Genet. 15, 133–141 (2014).
5. Carter, C. O. Genetics of common disorders. 87, 123–139 (2014).
6. Chakravarti, A. 2013 William Allan Award: my multifactorial journey. Am. J. Hum. Genet. 94, 325–333 (2014).
7. Emison, E. S. et al. Differential contributions of rare and common, coding and noncoding Ret mutations to multifactorial Hirschsprung disease liability. Am. J. Hum. Genet. 87, 60–74 (2010).
8. Arikkath, J. et al. β-Catenin regulates spine and synapse morphogenesis and function in hippocampal neurons during development. J. Neurosci. 29, 5435–5442 (2009).
9. Bareiss, S., Kim, K. & Lu, Q. β-Catenin/NPRAP (neural plakophilin-related armadillo repeat protein) interacts with and activates sphenogsin kinase 1. Biochem. J. 382, 717–723 (2004).
10. Martinez, M. C., Ochiiishi, T., Majewski, M. & Kosik, K. S. Dual regulation of neuronal morphogenesis by a β-catenin-cortactin complex and Rho. J. Cell Biol. 162, 99–111 (2003).
11. Silverman, J. B. et al. Synaptic anchorage of AMPA receptors by cadherins through neural plakophilin-related armadillo repeat protein receptor-binding protein complexes. J. Neurosci. 27, 8505–8516 (2007).
12. Laura, R. P. et al. The Erbin PDZ domain binds with high affinity and specificity to the carboxyl termini of β-catenin and ARVCF. J. Biol. Chem. 277, 12906–12914 (2002).
13. Kim, K. et al. Dendrite-like process formation and cytoskeletal remodeling regulated by β-catenin expression. Exp. Cell Res. 275, 171–184 (2002).
14. Daniel, J. M. & Reynolds, A. B. The catenin p120(catenin) interacts with Kaiso, a novel BTB/PDZ domain zinc finger transcription factor. Mol. Cell. Biol. 19, 3614–3623 (1999).
15. Yoon, H. G., Chan, D. W., Reynolds, A. B., Qin, J. & Wong, J. N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein. Mol. Cell. 12, 723–734 (2003).
16. Brigidi, G. S. & Palmitoylation of β-catenin by DHHC5 mediates activity-induced synapse plasticity. Nature Neurosci. 17, 522–532 (2014).
17. De Rubisi, S. et al. CYFIP1 coordinates mRNA translation and cytoskeleton remodeling to ensure proper dendritic spine formation. Neurology 79, 1169–1182 (2012).
18. Matsuoka, R. L. et al. Class 5 transmembrane semaphorins control selective mammalian retinal lamination and function. Neuron 71, 460–473 (2011).
19. We, W. Z. et al. Analysis of de novo copy number variations in a family affected with autism spectrum disorders using high-resolution array-based comparative genomic hybridization. Chin. J. Med. Genet. 29, 266–269 (2012).
20. Gardner, C. et al. A variant Cri du Chat phenotype and autism spectrum disorder in a subject with de novo cryptic microdeletions involving 5p15.2 and 3p24.3-25 detected using whole genomic array CGH. Clin. Genet. 67, 341–351 (2005).
21. Aoki, K. et al. Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. 82, 477–488 (2008).
22. Qiao, Y. et al. Phenotypic determinants of genomic variation in autism spectrum disorders. J. Med. Genet. 46, 680–688 (2009).
23. Sanders, S. J. et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 70, 863–885 (2011).
24. Gai, X. et al. Rare structural variation of synapse and neurotransmitter genes in autism. Mol. Psychiatry 17, 402–411 (2012).

Supplementary Information
is available in the online version of the paper.

Acknowledgements
We acknowledge the participation of all of the families in the AGRE, NIMH and SSC studies that have been a model of public participatory research. The AGRE is a program of Autism Speaks and is supported, in part, by grant U10NS058180 from the National Institute of Mental Health. The SSC used here was developed by the following principal investigators: A. Beaudet, R. Bernier, J. Constantino, E. Cook, F. Fonbonne, D. Geschwind, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, B. Peterson, J. Piggot, C. Nance, F. Stuedi, C. Walsh, E. Woznica, and W. Zoghbi. We thank Brain Atlas, C. et al. A variant Cri du Chat phenotype and autism spectrum disorder in a subject with de novo cryptic microdeletions involving 5p15.2 and 3p24.3-25 detected using whole genomic array CGH. Clin. Genet. 67, 341–351 (2005).

We thank V. Kustanovich (AGRE) for helping with access to Autism Diagnostic Observation Schedule severity data. D. Arking for sharing DNA from the SSC for TaqMan analysis. S. Maragkakis for developing and maintaining the manuscript (all authors); examined phenotype data for the female autism patients and requests for materials should be addressed to A.C. (aravinda@jhmi.edu).
METHODS
Human subjects and animal experimentation permissions. We studied 13 unrelated females, 12 from FEMs and one from a family with an affected girl and boy, from the AGRE6 and the NIMH collections (https://www.nimhgenetics.org/). Of these, 11 were of European, and one each of Hispanic and Native Hawaiian or Pacific Island ancestry. Studies of human subjects were approved by the Johns Hopkins Medicine Institutional Review Board (IRB NA_00015748). All protocols for animal care, use and euthanasia were reviewed and approved by the Institutional Animal Care and Use Committees of Johns Hopkins University (protocol MO12M412) and Duke University (protocol A229-12-08), and were in accordance with the Association for Assessment and Accreditation of Laboratory Animal Care guidelines. DNA sequencing, MEC2P sequencing. Each of the 13 individuals with autism was assessed for the four coding exons of MEC2P by PCR amplification of each exon and Sanger sequencing, performed at Beckman Coulter Genomics. The sequence traces were analysed in Sequencher version 4.7.

Exome sequencing and read mapping. We analysed ten female autism cases and one HapMap sample (NA18507) using the Agilent SureSelect Whole Exome capture (38 Mb) and SOLID 3 + 4 technologies. For three additional female individuals with autism, sequencing used the Illumina TruSeq Whole Exome capture (62 Mb) and Illumina technology; all Illumina exome experimental steps were performed at the Illumina sequencing centre. The SOLID and Illumina data were mapped to the human genome build 37 using the BEAST61 and BWA51 programs, respectively. Subsequently, the SAM output was converted to BAM output, duplicated to the human genome build 37 using the BFAST50 and BWA51 programs, respectively. For one HapMap sample (NA18507) using the Agilent SureSelect Whole Exome capture (62 Mb) and Illumina technology; all Illumina exome experimental steps were performed at the Illumina sequencing centre. The SOLID and Illumina data were mapped to the human genome build 37 using the BEAST61 and BWA51 programs, respectively. Subsequently, the SAM output was converted to BAM output, duplicated to the human genome build 37 using the BFAST50 and BWA51 programs, respectively. For one HapMap sample (NA18507) using the Agilent SureSelect Whole Exome capture (62 Mb) and Illumina technology; all Illumina exome experimental steps were performed at the Illumina sequencing centre. The SOLID and Illumina data were mapped to the human genome build 37 using the BEAST61 and BWA51 programs, respectively. Subsequently, the SAM output was converted to BAM output, duplicated to the human genome build 37 using the BFAST50 and BWA51 programs, respectively. For one HapMap sample (NA18507) using the Agilent SureSelect Whole Exome capture (62 Mb) and Illumina technology; all Illumina exome experimental steps were performed at the Illumina sequencing centre. The SOLID and Illumina data were mapped to the human genome build 37 using the BEAST61 and BWA51 programs, respectively. Subsequently, the SAM output was converted to BAM output, duplicated to the human genome build 37 using the BFAST50 and BWA51 programs, respectively.

Statistical analysis. Exome sequence. This study focused on variants of interest, defined as those that were absent in both dbsnp129 and 1000 Genomes low-pass sequencing data and were probably functionally deleterious (missense at residues conserved to zebrafish (human, chimp, dog, cow, mouse, rat and zebrafish from the University of California, Santa Cruz (UCSC) 46-way alignment), nonsense and canonical splice site changes). We compared each gene and its number of variants of interest with that expected on the basis of 10,000 replications of random sampling of 13 exomes from 71 female European controls. Genes having two or more variants of interest only in autism exomes were considered to be relevant candidates.

Allen Brain Atlas Data. The Allen Brain Atlas Microarray (Affymetrix Human Exon 1.0 ST data microarray summarized to genes (n = 17,630 genes) data set for the Developing Human Brain (8 weeks after conception to 40 years) was downloaded from the Allen Brain Atlas website on 24 February 2012. Linear regression was performed on the data set for age and brain region (in R software). Pearsonian correlations were calculated for each gene (X) and CTNND2 (Y); genes with absolute values greater than 0.3 were retained, corresponding to an experiment-wise P = 0.05 (17,630 comparisons) significance level. Pathway analyses were performed using MetaMx (https://www.metamx.org) and Ingenuity (Qiagen). Pathway analysis was performed on all correlated genes using the following categories: GOSTERMs (biological process, cellular compartment and molecular function) for function and UCSC TFBs. Functional assays. Generation of human CTNND2 and mouse Ctnnd2 constructs. Human CTNND2 was initially cloned into the pDONR221 Gateway vector. Subsequently, the human DNA was cloned into the pCS2 vector for zebrafish assays and pcDNA 6.2-N-EmgFP’-DEST vector (Gateway) for the neuronal assays.

Zebrafish gastrulation assays. Using a splice-blocking morpholino targeting zebrafish ctnnd2, one- to eight-cell stage embryos were injected (N = 50–180) and live embryos at the eight- to ten-stage somite stage were analysed for gastrulation phenotype including shortened body axes, longer somites, and broad and kinked notochords in morphant embryos. Embryos with phenotypes were then classified as class I or II depending on their severity (features of the convergence/extension phenotype include a shortened body axis, wider somites and a kinked notochord, with class I having one or two and class II having all three of these components, respectively). Specificity of the morpholino reagent was tested by co-injection of wild-type human CTNND2 mRNA. To test CTNND2 variants, injection cocktails containing morpholino and mutant human CTNND2 variants were injected and compared with the rescue condition of wild-type human CTNND2.

Zebrafish chordin expression assay. Zebrafish embryos were harvested at 90% epiboly stage and fixed in 4% paraformaldehyde at 4°C. Whole-mount RNA in situ hybridization was performed with a digoxigenin-labelled anti-chordin antibody probe synthesized by in vitro transcription (Roche). The chordin expression domain was measured in lateral view (‘L’ in Extended Data Fig. 7b). The middle point of the expression domain length and the centre of the embryo was linked with a dashed line (Extended Data Fig. 7b), along which the width of chordin expression domain (‘W’ in Extended Data Fig. 7b) was measured. Width:length ratio was calculated to quantify ectopic expression.

Immunoblotting. Cells were transfected with CTNND2 and Ctnnd2 epi expression constructs and harvested 48 h later. Protein lysates were immunoprecipitated using an anti-GFP antibody (Roche 1181446001) and immunoblotted with an anti-Flag antibody (Sigma F7425).

Neuronal cultures and synapse analysis. Hippocampi from day E18 rats or E17 mouse embryos were prepared and maintained as described elsewhere. At DIV8, the cells were transfected with GFP constructs (pcDNA6.2-N-EmgFP’-DEST) alone or fused to wild-type or variant allele containing CTNND2) and 500 ng of pcAG-DSRed2 using Lipofectamine 2000 (Life Technologies). On DIV16, cells were fixed with 4% paraformaldehyde and permeabilized, followed by immunostaining with primary antibodies against the appropriate target as described and their respective secondary antibodies. Neurons were imaged either on a Zeiss 510 confocal for standard analysis or on Zeiss 780/880 confocal microscopes for synaptic analysis, and analysed using ImageJ. Synapses were defined as puncta with overlapping signal between vGlut1 (Millipore catalogue number AB5905) and PSD95 (K28/43 clone from Neuronab) or vGlut1 and GluA (polyclonal antibody raised in rabbit against the C terminus of GluA subunit). To assess the expression of transfected CTNND2

©2015 Macmillan Publishers Limited. All rights reserved.
and its mutant alleles in the δ-catenin null background, we selected five pre-defined areas of interest with constant area in each dendrite (Extended Data Fig. 10).

Gene expression assays. Expression of ctnnd2a and ctnnd2b in zebrafish. To assess mRNA expression of the two zebrafish orthologues (ctnnd2a, ctnnd2b) of human δ-catenin, we performed PCR on normalized cDNA libraries (a gift of S. Maragh) from zebrafish at various developmental stages (50% epiboly, 75% epiboly, bud, 13 somite, 24 h after fertilization, and 3 days).

CTNND2 expression analysis in human and mouse tissues. To examine expression of CTNND2 in different human tissues, the Human MTC cDNA Panel 1 (Clontech catalogue number 636742, lot number 7080213), Human MTC cDNA Panel II (Clontech catalogue number 636743, lot number 6040176) and the Human Fetal MTC cDNA Panel (catalogue number 636747, lot number 5090557) were analysed by a TaqMan gene expression assay (catalogue number Hs00181643_m1) for CTNND2 and for a pipetting control (GAPDH, catalogue number 4333764T). Each tissue was tested in triplicate. Subsequently, the threshold cycle (Ct) values were averaged and the ΔCt values calculated between all of the tissues and the adult brain. The fold difference from brain was calculated as (1/(2ΔCt)) for each tissue.

49. Geschwind, D. H. et al. The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. Am. J. Hum. Genet. 69, 463–466 (2001).
50. Homer, N., Merriman, B. & Nelson, S. F. BFAST: an alignment tool for large scale genome resequencing. PLoS ONE 4, e7767 (2009).
51. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26, 589–595 (2010).
52. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010).
53. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164 (2010).
54. Brewer, G. J. Serum-free B27/neurobasal medium supports differentiated growth of neurons from the striatum, substantia nigra, septum, cerebral cortex, cerebellum, and dentate gyrus. J. Neurosci. Res. 42, 674–683 (1995).
Extended Data Figure 1 | Exome sequencing workflow and quality control. a, Workflow of exome analysis in this study. b, Variant recalibration metrics exhibiting why a 99% cutoff was used for truth sensitivity. c, Variant recalibration specificity versus sensitivity.
Extended Data Figure 2 | Sanger sequencing chromatograms. a, G34S variant in this study; b, R713C variants in this study; c, G34S in NA19020.
Extended Data Figure 3 | Principal component analysis of 6,211 shared autosomal single nucleotide polymorphisms in CEU, YRI, CHB/JPT, autism and NA19020 samples.
Extended Data Figure 4 | Read and amplicon metrics in CTNNB2 sequencing. a, Histogram of reads per sample. b, Average quality scores across the read across all samples with each sample represented by a separate line. c, Boxplot of coverage per amplicon.
Extended Data Figure 5 | Validation of deletions. a, In AU066818; b, in AU075604; c, in AU1178301 and AU1178202; d, in AU051503.
Neurodevelopmental Disease Phenotypes with Any Dosage Imbalance

Extended Data Figure 6 | CTNND2 CNVs from patients with neurodevelopment disorders studied using methods published in ref. 11.
Extended Data Figure 7 | Wnt defects in ctnnd2 zebrafish morphant embryos. a, Relative axin2 mRNA level in the ten-somite stage in control versus morphant embryos. b, Wholemount RNA in situ hybridization of chordin. Dorsal view in upper panels with the anterior aspect at the apex. The dorsal axis is marked with a red dashed line, and regions with high expression are marked (arrows) in control embryos. Lateral view in lower panels, length (L) and width (W) of chordin expression domains were measured. c, Quantification of chordin expression domains (length:width ratio) in injected embryos. d, Immunoblot showing a macromolecular interaction between Flag-tagged CTNNB1 and GFP-tagged CTNND2 with the corresponding variants. Two-sided t-tests were conducted: *P < 0.05, **P < 0.01 and ***P < 0.001, respectively. Sample size (n) is marked for each condition.
Extended Data Figure 8 | Functional in vitro modelling of δ-catenin missense variants in embryonic rat hippocampal neurons. a, Representation of spines along the dendrite in control and overexpression GFP vectors (empty or fused with wild-type or variant allele containing CTNND2 (G34S, R713C, A482T (control)). Cell counts for each construct were as follows: GFP (N = 32), GFP–WT (N = 27), GFP–G34S (N = 29), GFP–R713C (N = 26) and GFP–A482T (N = 29). b, Quantification of dendritic spine numbers and statistical comparisons by Tukey’s honestly significant test following ANOVA. Both * and ** indicate P < 0.05 than GFP and significantly different from wild type, respectively.
Extended Data Figure 9 | Gene expression of CTNND2 and co-expression with known autism genes.  

a. Expression of CTNND2 in various human fetal and adult tissues, shown as fold difference relative to adult brain.
b. RNA-Seq-based CTNND2 gene expression in the developing human brain (http://www.brainspan.org); shown are log2(RPKM expression) values at time-points from 8 weeks after conception to 40 years of age, with the lowest to highest expression coloured from navy blue to red. Controls for high expression, low to no expression and known autism genes are GAPDH, CFTR and MECP2, respectively.
Extended Data Figure 10 | Analysis of overexpression of transiently transfected neurons. Representation of average intensity of five individual regions of interest from a selected dendritic region. Quantitative comparison does not reveal a significant difference in expression levels of different variants of *CTNND2*. 