Early postnatal brain overgrowth and gene expression changes prefigure functional over-connectivity of the cortex in *Chd8* haploinsufficient mice

Philipp Suetterlin[1*], Shaun Hurley[1*], Conor Mohan[1*], Kimberley L. H. Riegman[1*], Marco Pagani[2], Angela Caruso[3,4], Jacob Ellegood[5], Alberto Galbusera[2], Ivan Crespo-Enriquez[1], Caterina Michetti[3,6], Robert Ellingford[1], Olivier Brock[7], Alessio Delogu[7], Philippa Francis-West[1], Jason P. Lerch[5], Maria Luisa Scattoni[3], Alessandro Gozzi[2], Cathy Fernandes[8,9], & M. Albert Basson[1,9^]

[1]Department of Craniofacial Development and Stem Cell Biology, King's College London, Floor 27, Guy's Hospital Tower Wing, London, SE1 9RT, UK
[2]Functional Neuroimaging Laboratory, Center for Neuroscience and Cognitive Systems @ UniTn, 38068 Rovereto, TN, Italy
[3]Neurotoxicology and Neuroendocrinology Section, Department of Cell Biology and Neuroscience, Istituto Superiore di Sanità, Viale Regina Elena 299,00161 Rome, Italy
[4]School of Behavioural Neuroscience, Department of Psychology, Sapienza University of Rome, via dei Marsi 78, 00185, Rome, Italy
[5]Department of Medical Biophysics, University of Toronto, Mouse Imaging Centre, Hospital for Sick Children, 25 Orde Street, Toronto, Ontario M5T 3H7, Canada
[6]Center for Synaptic Neuroscience and Technology, Istituto Italiano di Tecnologia, Genova, Italy
[7]Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[8]MRC Social, Genetic & Developmental Psychiatry Centre, PO82, Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, London SE5 8AF, UK
[9]MRC Centre for Neurodevelopmental Disorders, King's College London, 4th floor, New Hunt's House, London SE1 1UL, UK

* First co-authors
^Corresponding author

Running title: Abnormal brain development in *Chd8*$^{+/-}$ mice

Pages: 60
Figures: 6
Abstract: 180 words
Introduction: 643 words
Discussion: 1494 words

ABSTRACT

Truncating *CHD8* mutations are amongst the highest confidence risk factors for autism spectrum disorders (ASD) identified to date. Here, we report that *Chd8* heterozygous mice display subtle brain hyperplasia shortly after birth, hypertelorism, early motor delay, pronounced hypoactivity and anomalous responses to social stimuli. Whereas gene expression in the neocortex is only mildly affected at mid-gestation, over 600 genes are differentially expressed in the early postnatal neocortex. Genes involved in cell adhesion and axon guidance are particularly prominent amongst the down-regulated transcripts. Resting-state functional MRI identified increased synchronised activity in cortico-hippocampal and auditory-parietal networks in *Chd8* heterozygous mutant mice, implicating altered connectivity as a potential mechanism underlying the behavioural phenotypes. Together, these data suggest that altered brain growth and diminished expression of important neurodevelopmental genes that regulate long-range brain wiring result in distinctive anomalies in functional brain connectivity in $Chd8^{+/-}$ mice. Human imaging studies have consistently found evidence for changes in functional connectivity in ASD cohorts, most commonly long-range under-connectivity. Our data suggest that *CHD8* haploinsufficiency represents a specific subtype of ASD where neuropsychiatric symptoms are underpinned by long-range over-connectivity.

SIGNIFICANCE STATEMENT

De novo mutations in the chromatin remodeling factor CHD8 cause ASD with high penetrance, making *CHD8* one of the leading ASD candidate genes. We established a *Chd8* heterozygous mouse model that recapitulates anatomical phenotypes seen in patients with *CHD8* haploinsufficiency and displays a unique complement of behavioural phenotypes. Increased functional connectivity observed in cortical and hippocampal areas suggests that neuropsychiatric phenotypes associated with CHD8 deficiency may be the result of distinct connectivity changes. Transcriptomic analyses highlighted dysregulation of axon guidance genes as a possible underlying mechanism. Together, our data suggest that *CHD8* haploinsufficiency represents a distinct ASD subtype characterised by a unique set of anomalous behaviours, including increased responsiveness to social stimuli, and functional connectivity changes in mice.

INTRODUCTION

Autism spectrum disorder (ASD) is diagnosed on the basis of socio-communicative deficits and repetitive, perseverative behaviours with restricted interests (APA, 2013). ASD is frequently associated with comorbidities like hyper-sensitivity to sensory stimuli, seizures and anxiety (Croen et al., 2015; Jeste and Tuchman, 2015; Tavassoli et al., 2014). The phenotypic and genetic heterogeneity of ASD has hampered the elucidation of the molecular mechanisms that may underlie specific aberrant behaviours. However, the recent identification of de novo, likely gene disrupting (LGD) mutations that show highly significant associations with autism (Iossifov et al., 2014; Neale et al., 2012; O'Roak et al., 2014; O'Roak et al., 2012b; Talkowski et al., 2012) provides an opportunity to phenotype and molecularly characterise genetically defined ASD subtypes.

Exome sequencing studies of several thousand simplex families detected de novo, likely gene disrupting (LGD) mutations in the *CHD8* (Chromodomain Helicase DNA binding factor 8) gene (Iossifov et al., 2014; Neale et al., 2012; O'Roak et al., 2014; O'Roak et al., 2012b; Talkowski et al., 2012). Patients with *CHD8* mutations are characterised by a high prevalence of autism, macrocephaly, facial dysmorphisms, motor delay and hypotonia, intellectual disability and gastro-intestinal problems, and less commonly by anxiety and seizures (Bernier et al., 2014; Merner et al., 2016; Stessman et al., 2017; Stolerman et al., 2016). *CHD8* encodes an ATP-dependent chromatin remodelling protein of the chromodomain helicase DNA binding family (Thompson et al., 2008; Yuan et al., 2007). The recruitment of CHD8 to gene promoters in mouse and human neural progenitors is strongly associated with their active transcription, while *CHD8* knock-down in these cells results in downregulation of a plethora of ASD-associated genes (Cotney et al., 2015; Sugathan et al., 2014).

Altered brain connectivity, characterised by local over-connectivity and long-range under-connectivity, has been hypothesised to underpin some of the neuropsychiatric phenotypes observed in ASD (Belmonte et al., 2004; Just et al., 2004). Resting-state functional MRI (rsfMRI) studies in ASD patients have provided evidence for reduced long-range synchronisation in spontaneous brain activity (reviewed in Picci et al., 2016). Increased long-range connectivity has also been reported in a subset of cases (Di Martino et al., 2014), consistent with the phenotypic heterogeneity of ASD. Thus, the exact nature of aberrant functional connectivity in ASD may depend on the specific underlying aetiology.

Few rsfMRI studies have been performed in ASD mouse models (Liska and Gozzi, 2016). As one example, homozygous *Cntnap2* mouse mutants exhibit hypo-connectivity of the default mode network, a network that consists of the synchronised spontaneous activation of the frontal, cingulate and retrosplenial cortex (Liska et al., 2017). Reduced functional connectivity in the default mode network is consistently detected in ASD patients in human rsfMRI studies (Cherkassky et al., 2006) and these findings recapitulate analogous clinical observations in humans with *CNTNAP2* mutations (Scott-Van Zeeland et al., 2010).

Three groups recently reported analyses of *Chd8*$^{+/-}$ mouse models (Gompers et al., 2017; Katayama et al., 2016; Platt et al., 2017). They consistently found megalencephaly, subtle but wide-spread transcriptional changes and behavioural anomalies across the three models. Katayama et al. identified a delay in embryonic neural development caused by de-repression of the RE1-silencing transcription factor (REST) and attenuated expression of its target genes. Gompers et al. detected aberrations in RNA splicing that may contribute to transcriptional dysregulation. Platt et al. found evidence for electrophysiological changes in striatal circuits. No differences were seen in overall white mater organisation by diffusion

tensor imaging suggesting that major axon tracts are intact in the *Chd8*$^{+/-}$ brain (Gompers et al., 2017).

To further understand how *Chd8* haploinsufficiency disrupts brain development and contributes to ASD we generated a *Chd8*$^{+/-}$ mouse model where exon 3 of *Chd8* is deleted, equivalent to early nonsense and frameshift mutations found in patients (Barnard et al., 2015). We characterised brain growth and gene expression changes during gestation and early postnatal development and employed rsfMRI to probe functionally connected brain networks in these mutants.

RESULTS

A mouse line with a conditional *Chd8* allele was produced through homologous recombination in C57Bl/6J embryonic stem cells (Fig. 1A, B). *Chd8*^{flox} mice were crossed with the ubiquitously expressing *βactin*-Cre line (Lewandoski and Martin, 1997) on a C57Bl/6J background to generate *Chd8*$^{+/-}$ mice (Fig. 1C). Cre-mediated deletion of *loxP*-flanked (flox) exon 3 results in an early frameshift and termination of translation at amino acid 419, predicted to produce a protein that lacks all functional domains, equivalent to nonsense and frameshift mutations terminating CHD8 at amino acids 62 and 747 in patients (Barnard et al., 2015).

Quantitative RT-PCR (qRT-PCR) on RNA isolated from E12.5 and P5 neocortices using primers spanning the exon 3/4 boundary found that total *Chd8* expression was reduced by 64% (p=0.006, student's t-test) and 52% (p=0.01, student's t-test), respectively (Fig. 1D). Western blot experiments using an antibody against the N-terminal region of CHD8 (amino acids 325 – 350) on lysates from E12.5 neocortex revealed a 51% reduction in full-length CHD8 protein levels in heterozygotes compared to controls (Fig. 1E,F), validating our *Chd8*$^{+/-}$ mice as a suitable model for *CHD8* haploinsufficiency. Importantly, we found no evidence for a truncated protein product of 419aa (~45kDa) that may have resulted from translation of any mutant transcript (Fig. 1E). Indeed, repeating the qRT-PCR analysis at E12.5 with primers spanning the exon 1/2 boundary (upstream of the recombination event) reveals a reduction in *Chd8* expression by 52% (Fig. 1G), indicating that the mutant transcript is most likely subject to nonsense-mediated decay.

Chd8 heterozygous mice have specific craniofacial and structural brain phenotypes

Humans with truncating mutations in a single *CHD8* allele often present with macrocephaly (64%) and distinct craniofacial phenotypes (89%), which include hypertelorism (wide-set eyes, 67%) (Bernier et al., 2014; Stessman et al., 2017). We characterised the cranioskeleton of *Chd8*$^{+/-}$ mice by µCT to ask whether these phenotypes are also present in *Chd8*$^{+/-}$ mice. The interorbital distance (landmarks 8-9, Fig. 2C,D) is significantly wider in *Chd8*$^{+/-}$ mice compared to controls, indicative of a hyperteloric phenotype (*p=0.0273, student's t-test; Fig. 2C,D,F). In addition, the anterior-posterior length of the interparietal bone (landmarks 4-5) is increased in *Chd8*$^{+/-}$ animals (**p=0.0025, student's t-test; Fig. 2A,B,E), suggestive of more wide-spread craniofacial anomalies associated with *CHD8* haploinsufficiency.

To examine whether structural brain abnormalities are present in *Chd8*$^{+/-}$ mice, their brains were compared to *Chd8*$^{+/+}$ littermates by high resolution MRI (Fig 2G). Total brain volume was increased by 2.7% in *Chd8*$^{+/-}$ mice (476mm^3 vs. 463mm^3, p=0.048, FDR=15%, Fig. 2H). Accordingly, several brain regions, including cortical areas, hippocampus and parts of the cerebellum showed volumetric increases (Fig. 2G, H, Table 2-1). Structural alterations of these brain areas have been implicated in autism (Blatt, 2012; Donovan and Basson, 2017; Ecker, 2016) providing potential neural substrates for the autism phenotype associated with *CHD8* haploinsufficiency in humans.

Chd8$^{+/-}$ mice show abnormal activity levels and differences in social interaction

CHD8 heterozygosity is associated with autism in the human population (Iossifov et al., 2014; Neale et al., 2012; O'Roak et al., 2014; O'Roak et al., 2012b; Talkowski et al., 2012). We therefore asked whether *Chd8*$^{+/-}$ mice exhibited any signs of socio-communicative deficits, repetitive behaviours or cognitive inflexibility.

Chd8 heterozygous pups displayed signs of delayed motor development in the first two weeks after birth. *Chd8*$^{+/-}$ pups took slightly longer than wildtype littermates to develop an effective righting reflex over time (*p=0.014, one-way repeated-measures ANOVA; Fig. 3A). Correspondingly, *Chd8*$^{+/-}$ pups spent more time engaged in unsuccessful attempts to turn over on their stomachs as measured during the spontaneous motor behaviour observations during USV recordings (P6: *p=0.0312, P8: *p=0.0354, student's t-test; Fig. 3B). Once they were able to move around the cage, mutant pups spent on average more time in locomotion than wildtype littermates suggestive of hyperactivity (**p=0.009, student's t-test; Fig. 3C).

In the three-chamber sociability test, adult *Chd8*$^{+/-}$ mice spent significantly more time in the chamber with the novel age- and sex-matched conspecific mouse than in the other chambers, thus demonstrating normal sociability (Fig. 3D). Interestingly, rather than displaying sociability deficits, mutant mice spent slightly but significantly more time in the chamber containing the mouse, compared to controls (*p=0.029, between subjects ANOVA; Fig. 3D). *Chd8*$^{+/-}$ mice also spent more time investigating conspecific mice in a reciprocal social interaction test (*p=0.015, between-subjects ANOVA; Fig. 3E). A quantitative olfactory habituation/dishabituation test revealed an increased interest in an odour with social significance (urine) in *Chd8*$^{+/-}$ mice compared to controls, as measured by the cumulative time spent sniffing the odour (*p=0.03, ***p=0.0002, student's t-test; Fig. 3F). No difference in the time spent investigating a non-social (banana) odour was observed, implying an increased interest specifically in social cues and an otherwise normal capacity for odour discrimination (Fig. 3F).

Examination of these animals in the open field arena revealed a marked hypo-activity in *Chd8*$^{+/-}$ mice (***p=10^{-9}, between subjects ANOVA; Fig. 3G,H). The hypo-active phenotype was also present in mutant mice in their homecage environment by measuring activity on a running wheel over a one-week period (*p=0.019, repeated measures ANOVA; Fig. 3I). The open field test did not show any evidence of anxiety in these mice, i.e. an increased reluctance to enter the inner, most exposed area of an open field arena (Fig. 3J). This was confirmed in the light/dark box test that showed no difference between wildtype and mutant mice (Fig. 3K). Forelimb grip strength was slightly but significantly reduced in mutant mice (*p=0.045 (males) *p=0.042 (females),between subjects ANOVA; Fig 3L) but *Chd8*$^{+/-}$ mice showed normal abilities on the revolving rotarod, indicating that a reduced capacity to perform motor tasks was unlikely to be the cause of the hypo-active phenotype (Fig. 3M). No evidence of repetitive behaviours was observed by assessing marble burying and self-grooming behaviours (Fig. 3N,O). In fact, mutants showed slightly delayed marble burying behaviour, most likely due to their general hypoactivity (*p=0.04, ***p=0.0004, student's t-test; Fig. 3N).

Spatial learning abilities and cognitive flexibility were assessed in the hippocampus-dependent Morris water maze test. *Chd8*$^{+/-}$ mice performed normally in the learning part of this test (Fig. 3P). In a reversal paradigm, these mice were also indistinguishable from wildtype littermates, implying normal cognitive, spatial learning abilities and flexibility (Fig. 3P). Finally, no differences in the number of ultrasonic vocalisations (USVs) of pups separated from the nest were recorded, indicating no obvious communication deficits (Fig. 3Q).

In summary, adult *Chd8*$^{+/-}$ mice exhibited a hypo-active phenotype, while pups showed evidence for hyperactivity and delayed motor development. *Chd8*$^{+/-}$ mice showed no evidence for perseverative and repetitive behaviours. Intriguingly, a heightened interest in social cues was observed in *Chd8*$^{+/-}$ mice.

Chd8 haploinsufficiency causes general growth delay but postnatal brain overgrowth

To characterise the developmental delay observed during behavioural testing and the brain overgrowth phenotype detected by MRI analysis further, we followed postnatal body weight trajectories and correlated these with brain weights at different developmental stages for wildtype and *Chd8* heterozygous pups. A survey of body weight during postnatal development identified significant growth retardation of *Chd8*$^{+/-}$ pups from postnatal day 4 onwards, a phenotype replicated in an independent cohort, that culminated in a highly significant statistical difference in body weight between the two groups at P35 (15.8%, ***p<0.0001, student's t-test; Fig. 4A). Brain and body weight were well correlated in both wildtype and heterozygous mice (r^2=0.25, p=0.0004 and r^2=0.28, p=0.005, respectively). *Chd8* mutants had higher brain weights compared to their wildtype littermate controls with equivalent body weight (Fig. 3B). A group-wise comparison confirmed the significant increase in normalised brain weight in *Chd8*$^{+/-}$ mice compared to wildtype littermates at P35 (20.4% increase, ***p<0.0001, student's t-test, Fig. 4C).

To identify the developmental time window when signs of brain overgrowth first emerge, normalised brain weights were examined at earlier developmental stages. At P7, normalised brain weights were already significantly larger in *Chd8*$^{+/-}$ pups compared to wildtype littermate controls (9.3%, ***p=0.0009, student's t-test) with more subtle differences between the groups observed at P0 (6.7%, *p=0.01, student's t-test), suggesting that the

phenotype emerged in the perinatal period. Indeed, the same analysis at E16.5 showed a trend towards increased brain size but did not detect any significant differences in *Chd8*$^{+/-}$ embryos compared to wildtype littermates (4.3%, p=0.067, student's t-test; Fig. 4D). Together, these analyses suggested that small differences in brain growth over several days during late embryonic and early postnatal development may be responsible for small, progressive increases in brain size. Indeed, we did not detect any significant differences in cortical ventricular zone (VZ) proliferation as measured by phospho-histone H3 immunostaining at E12.5, E16.5 and P0 (Fig. 4E,F). However, subtle or short-lived increases in progenitor proliferation in the VZ cannot be completely ruled out.

In summary, larger brain size in adult mice implicated an uncoupling of postnatal brain growth from overall growth as a potential cause of the macrocephalic phenotype observed in patients with *CHD8* mutations. Moreover, the developmental motor delay observed in *Chd8*$^{+/-}$ pups was underpinned by reduced body weight in the first few weeks after birth, a phenotype that resolved by 3 months of age, when body weights of *Chd8* heterozygous mice were indistinguishable from wildtype littermate controls (data not shown).

CHD8 controls the expression of ASD-associated axon guidance genes in the early postnatal neocortex.

To gain insights into the transcriptional programmes that may underlie the subtle brain overgrowth and abnormal behaviours observed in *Chd8*$^{+/-}$ mice, we performed RNA-seq analysis on dissected neocortical tissue at two stages: 1) At E12.5, when *Chd8* expression peaks (Durak et al., 2016) and neural progenitor cells predominate, and 2) At P5, when many

developmental processes with relevance for ASD aetiology, such as axon growth and guidance and synaptogenesis, are taking place.

Surprisingly, only 5 genes, including *Chd8*, showed significant (FDR<0.05) differential expression in *Chd8*$^{+/-}$ embryos at E12.5 (Fig. 5A, Table 5-1). By contrast, 649 genes (FDR<0.05) were differentially expressed in the P5 neocortex, with over two thirds of them downregulated (Fig. 5B, Table 5-2).

Comparing all differentially expressed genes (DEGs) from the P5 dataset with the SFARI autism gene list identified 56 shared genes, representing a highly significant enrichment of ASD-associated genes in the differentially expressed gene set ($p=1.06x10^{-10}$ (OR=2.87), Fisher's exact test for count data, Fig. 5C, Table 5-3). 53 out of 56 of these ASD-associated genes were down-regulated (95%, Table 5-3). We also overlapped our gene set with high confidence (SFARI categories 1&2) ASD candidates ($p=3.26x10^{-4}$ (OR=4.61), Fisher's exact test for count data Fig. 5D). Nine genes, representing 16% of all SFARI category 1 & 2 genes, were present in our differentially expressed gene set at P5. All of these high confidence ASD candidate genes were down-regulated (Table 5-2).

Amongst the upregulated gene set, the most significant KEGG pathways, molecular functions and biological processes were related to the ribosome and oxidative phosphorylation, whereas the downregulated gene set included categories related to cell adhesion, axonal guidance and calcium signaling pathways (Fig. 5E, Fig. 5-1A, Tables 5-3-5-9). Identification of potential regulatory transcription factors was performed using Enrichr, which found over-representation of Suz12 targets in the down-regulated gene set (Fig. 5-1B). Suz12 is a component of the Polycomb repressor complex 2 (PCR2) and is required for both histone

methyl transferase and gene silencing activities of PRC2 (Cao and Zhang, 2004). The observation that Suz12 targets are over-represented in the down-regulated gene set offers a potential mechanistic explanation for the down-regulation of some of the identified genes. None of the genes that code for components of PRC2, including *Suz12*, are differentially expressed at P5, suggesting that the enrichment seen was unlikely to be due to direct transcriptional dysregulation of polycomb gene expression at this stage.

The significant enrichment of cell adhesion and axonal guidance genes in the down-regulated gene set at P5 suggested the possibility that long-range connectivity might be disrupted in *Chd8* heterozygous neocortices.

$Chd8^{+/-}$ mice exhibit over-connectivity in cortical and hippocampal networks.

Functional brain connectivity can be probed by rsfMRI, a method based on localised changes in blood flow upon neural activity. These changes are commonly measured as fluctuations in the blood-oxygen-level dependent (BOLD) signal, with functionally connected areas acting in synchrony. To test whether the observed gene expression changes at P5 indeed prefigured functional alterations in mature brain networks, we compared *Chd8* heterozygous and wildtype adult mice by rsfMRI. A regionally unbiased analysis for long-range connectivity changes revealed hotspots for increased connectivity in $Chd8^{+/-}$ mice compared to wildtype littermate controls, which included the entorhinal, retrosplenial, auditory cortical and posterior hippocampal areas (t-test, $p<0.05$ FEW cluster-corrected, with cluster-defining threshold $t_{24}>2.06$, $p<0.05$; orange areas in Fig. 6A). This analysis suggested that hyper-connected areas were predominantly located on the left side of the brain. A re-analysis of these results without the use of cluster correction revealed the presence of foci with increased connectivity also on the right side, mirroring the effects observed on the left (dark red areas

in Fig. 6A). Inter-hemispheric mapping of rsfMRI connectivity strength in previously characterized rsfMRI network systems of the mouse brain (Sforazzini et al., 2014), revealed increased cortical connectivity in auditory regions ($p<0.05$, Student t test, uncorrected), although the effect did not survive false discovery rate correction ($q=0.05$) for multiple comparison across the rsfMRI networks probed. We next used a seed-based approach to specifically probe regions with altered connectivity to these hotspots to reveal the brain networks affected. Most strikingly, this revealed a reciprocal increase in connectivity between ventral hippocampus and auditory cortical regions in *Chd8* mutant mice (t-test, $p<0.05$ FEW cluster-corrected, with cluster-defining threshold $t_{24}>2.06$, $p<0.05$; Fig. 6B,C). Seed placement in the auditory cortex further revealed increased connectivity of this region with both cingulate and entorhinal cortices (Fig. 6B), whereas a hippocampal seed uncovered strengthened long-range connectivity with somatosensory and visual cortices (Fig 6C). These findings suggested that specific cortical-hippocampal circuits involved in sensory processing may be altered in *Chd8*$^{+/-}$ mice.

DISCUSSION

CHD8 is one of the highest confidence ASD-associated genes to emerge from recent exome sequencing studies (Bernier et al., 2014; Iossifov et al., 2014; Neale et al., 2012; O'Roak et al., 2012a; Talkowski et al., 2012). We therefore expected *Chd8*$^{+/-}$ mice to present with robust, autism-associated behaviours. *Chd8*$^{+/-}$ mice displayed delayed motor development and distinctive behavioural anomalies that featured a heightened interest in social cues, but did not include repetitive and perseverative behaviours or communication deficits. Gene expression analysis showed little evidence for transcriptional dysregulation at mid-embryonic stages, but revealed disruption of key developmental processes involved in establishing brain connectivity in the early postnatal neocortex. Many of these dysregulated transcripts, which were predominantly down-regulated, are themselves ASD-associated genes, providing a strong rationale for the ASD phenotype associated with *Chd8* deficiency in patients. Significantly, the disruption of these neurodevelopmental transcriptional networks was succeeded by alterations in functional connectivity in the adult brain. Resting-state functional MRI analysis found evidence for over-connectivity between sensory regions in the neocortex and limbic cortical regions. Most notably, the auditory cortex showed a global increase in functional connectivity that involved connections to other cortical areas and reciprocal strengthening of connectivity to the ventral hippocampus. These findings identify a crucial developmental function for *CHD8* in the early postnatal period and, for the first time, associate *CHD8* haploinsufficiency with altered functional connectivity of brain areas with possible relevance to ASD. We conclude that *Chd8* haploinsufficiency results in a distinct complement of neurodevelopmental and behavioural anomalies in mice and propose that these mice will represent a useful model for elucidating the molecular mechanisms and circuit abnormalities underlying the ASD subtype associated with *CHD8* haploinsufficiency in humans.

Chd8$^{+/-}$ mice as a model for human *CHD8* haploinsufficiency syndrome

We report several phenotypes in *Chd8*$^{+/-}$ mice that have been associated with *CHD8* haploinsufficiency in humans. These include hypertelorism and delayed motor development. We also describe mild postnatal brain overgrowth, which may underlie the macrocephaly observed in the majority of patients described by Bernier et al. (Bernier et al., 2014). Together, these findings indicate that core functions of CHD8 in craniofacial and brain development are conserved between mouse and human, with previous studies suggesting that this is also true for non-mammalian species such as zebrafish (Bernier et al., 2014; Sugathan et al., 2014).

Autism is one of the most prevalent features observed in human *CHD8* haploinsufficiency, with a reported penetrance to date of 96% (Stessman et al., 2017). Most strikingly, *Chd8*$^{+/-}$ mice show a delay in early motor development, and there is a growing body of evidence suggesting that delayed motor milestones in toddlers predate and predict the emergence and severity of language deficits in later life (Bedford et al., 2016; Chinello et al., 2016). Of note, the two longitudinal case reports in the literature describe early motor delay in both patients with *CHD8* haploinsufficiency (Merner et al., 2016; Stolerman et al., 2016). We did not observe repetitive or perseverative behaviours in *Chd8*$^{+/-}$ mice. While possible group differences in our data set could have been confounded by the general hypoactivity of *Chd8*$^{+/-}$ mice, the overall conclusions are in agreement with other published reports (Gompers et al., 2017; Katayama et al., 2016; Platt et al., 2017). Finally, *Chd8*$^{+/-}$ mice show an apparent heightened interest in social cues, indicating that altering the *Chd8* gene dosage during development can impact socially motivated behaviours. An increased duration of contacts in the social investigation test was also seen in two other published behavioural analyses of

Chd8 heterozygous mouse models (Fig. 3E) (Katayama et al., 2016; Platt et al., 2017). In addition, these studies reported subtle social deficits, albeit on a background of unusually high anxiety in all test groups (Katayama et al., 2016), a known confound in behavioural tests. A key characteristic of autism is restricted behaviours or interests, which often manifest as hyper- or hypoactivity to sensory input or unusual interest in sensory stimuli, for example excessive smelling or touching of objects (Constantino and Charman, 2016). One may speculate that the excessive smelling of social cues and the increased duration of social contacts observed in our *Chd8*$^{+/-}$ mice may be indicative of behavioural abnormalities in these domains.

Dysregulation of the cortical transcriptome in *Chd8* heterozygous mice.

Chd8 levels peak at E12, and gradually decline during embryonic development and into postnatal stages (Durak et al., 2016). Interestingly, our data showed that *Chd8* heterozygosity has little effect on the neocortical transcriptome at E12.5, but widespread transcriptional dysregulation was apparent by early postnatal stages. One possible explanation for this finding could be that the higher absolute levels of *Chd8* at E12.5 are sufficient to maintain normal gene expression even in the heterozygous state, but as *Chd8* levels fall in late embryonic and early postnatal stages the gene dosage provided by a single functional copy of *Chd8* is no longer sufficient to maintain a normal transcriptomic programme. A second possibility is that *Chd8* heterozygosity causes subtle dysregulation of other regulatory genes early on, leading to knock-on effects that manifest in gene expression changes later in development. In support of this idea, we see over-representation of PRC2 binding motifs in the *Chd8*$^{+/-}$ down-regulated gene set at P5 in the absence of expression changes in PRC2 components themselves, thus suggesting secondary transcriptional effects. Similarly, Katayama and colleagues reported de-repression of REST in the *Chd8* heterozygous brain

leading to reduced expression of REST target genes at E14.5 (Katayama et al., 2016). Irrespective of the underlying mechanism, our data identify early postnatal development as a key stage at which transcriptional changes caused by *Chd8* heterozygosity may precipitate ASD-related phenotypes.

An expanding number of ASD risk genes have roles in axon guidance, synapse development and plasticity (Bourgeron, 2015). We detected significant enrichment of genes in these functional categories in our down-regulated gene set. These include the major Slit protein receptors *Robo1* and *Robo2* (Tables 5-2 and 5-5). Slit proteins are critical for establishing several major axonal tracts in the developing forebrain (Bagri et al., 2002). In addition, down-regulated genes in P5 heterozygous animals were enriched for cell adhesion molecules, including *L1cam*, which has important roles in neuron migration, neurite outgrowth, modulation of actin cytoarchitecture and axon targeting (Maness and Schachner, 2007). In support, Gompers et al. found enrichment for axon growth and guidance factors amongst down-regulated genes in their M3 module (Gompers et al., 2017). Interestingly, Sugathan and colleagues also found enrichment for genes associated with the GO terms 'cell adhesion', 'axon guidance' and 'neuron differentiation' in their down-regulated gene set. They conducted RNA-seq analysis on human iPSC-derived neural progenitors, thus suggesting that expression of these important developmental gene sets is regulated by CHD8 in both mouse and human cells (Sugathan et al., 2014). Taken together, these findings indicate that *Chd8* heterozygosity defines a transcriptional programme characterised by diminished expression of key neurodevelopmental regulators that are predicted to affect a complement of cellular functions essential for the appropriate wiring of the brain.

Increased functional connectivity in sensory networks.

RsfMRI revealed altered functional connectivity in several cortical networks in *Chd8*$^{+/-}$ mice. It seems likely that altered connectivity is the consequence of some of the disrupted brain wiring pathways uncovered by our RNA-seq experiments, but this hypothesis will require further in–depth scrutiny. More importantly, it will be critical to investigate whether these connectivity changes are pertinent to any of the behavioural anomalies in *Chd8* heterozygous mice or the ASD phenotype in patients with *CHD8* haploinsufficiency. Intriguingly, we see over-connectivity in networks involving the auditory cortex and the hippocampus. Auditory processing deficits in ASD are well documented and range from a lack of lateralisation to a general delay in network maturation (Bruneau et al., 1992; Edgar et al., 2015), although the functional behavioural consequences of these deficits are not clear. Furthermore, over-responsivity to sensory stimuli is frequently observed in ASD patients, can affect all sensory modalities and appears to be positively correlated with the severity of autistic traits (reviewed in Sinclair et al., 2017; Tavassoli et al., 2014). Although a definitive causal relationship is difficult to establish, it has been hypothesised that sensory over-responsivity may trigger compensatory and avoidance behaviours that promote the emergence of core behavioural autism traits (Marco et al., 2011). Whether this would be equally the case in ASD mouse models is an open question.

Summary and outlook.

Our data provide the first indication that long-range functional connectivity is altered in a *Chd8*$^{+/-}$ mouse model. A recent rsfMRI study involving over 150 male probands with an ASD diagnosis and nearly 200 typically developing individuals described over-connectivity between sensory cortices and subcortical structures as a central feature in ASD (Cerliani et al., 2015). It will be very important to determine whether abnormalities similar to those we describe in *Chd8*$^{+/-}$ mice are detected in patients with *CHD8* haploinsufficiency.

While we cannot at this stage establish a direct causal relationship between transcriptional, connectivity and behavioural phenotypes, our data point towards functionally relevant abnormalities in connectivity of sensory cortical areas in *Chd8* mutant mice that underpin behavioural phenotypes. Future studies will address these hypotheses.

METHODS

Chd8 gene targeting

A 14.84kb genomic DNA fragment was subcloned from C57BL/6 BAC clone (RP23: 318M20) into pSP72 (Promega). This fragment encompassed a long homology arm (LA) of 9.45kb 5' and a short homology arm (SA) of 4.4kb 3' of the site to be targeted. The targeting construct was generated by inserting a loxP/FRT-PGK-gb2-Neo cassette 214bp 3' of exon 3 (ingenious targeting laboratory (iTL), Ronkonkoma, NY, USA). An additional single loxP site containing a BclI restriction site for Southern blot screening was inserted 5' of exon 3. The final targeting construct of 18.8 kb was linearised by NotI digestion and electroporated into C57BL/6J ES cells. G418-resistent clones were selected, screened by PCR and Southern blot for successful homologous recombination. Five clones with successful recombination were identified (Figure S1) and two clones (124 and 254) were injected into Balb/c blastocysts (iTL). Resulting chimaeras were bred with Flpe deleter mice on a C57BL/6J background to excise the neo cassette and produce *Chd8*$^{flox/+}$ mice (Figure 1). *Chd8*$^{flox/+}$ mice were then crossed with *β-actinCre* mice (Lewandoski and Martin, 1997) to generate a *Chd8* null (*Chd8*$^{-}$) allele. *β–actinCre;Chd8*$^{+/-}$ mice were crossed with C57BL/6J mice to remove the Cre transgene and establish a *Chd8*$^{+/-}$ line.

Mice

Experimental mice were produced by *Chd8*$^{+/-}$ x C57BL/6J crosses, taking care to equalise paternal or maternal inheritance of the *Chd8* null allele, especially for behavioural experiments. For genotyping, genomic DNA was routinely extracted from ear or tail samples (or yolk sac for E12.5 embryos) using Proteinase K digestion or the HotSHOT method (Truett et al., 2000). Genotyping reactions were then performed for the presence of *Chd8* wildtype and null alleles using the following primer pair: FW: CCC ACA TCA AGT GGC

TGT AA, Rev: GGT AGG GAA GCA GTG TCC AG. Thermal cycles were as follows: 94°C, 5 minutes; 35X (94°C, 30sec; 58°C, 30sec; 72°C, 30sec); 72°C, 5 minutes. This yielded a PCR product of 395bp for the null allele and 1.1kb for the wildtype allele.

Western Blot

Telencephalic vesicles were dissected from E12.5 embryos and whole cell protein prepared by lysing in 8M urea, 1% CHAPS, 50mM Tris (pH 7.9) lysis buffer containing protease inhibitors (PMSF, Pepstatin A, Leupeptin, Aprotinin; Roche) and a phosphatase inhibitor cocktail (Sigma). Lysates were rotated at 4°C for 30 mins followed by DNA removal by centrifugation. Supernatant was transferred to a fresh tube and stored at -80°C. Protein loading samples were made by diluting samples in Laemmli buffer containing 10% β-mercaptoethanol, followed by boiling at 95°C for 5 minutes.

Samples were loaded (10µg total protein per lane) onto a Mini-PROTEAN pre-cast gel (Bio-Rad) and resolved using gel electrophoresis. Protein was transferred to a nitrocellulose membrane (Bio-Rad) which was then blocked in 5% non-fat milk powder (Bio-Rad) and 1% bovine serum albumin (BSA, Sigma) in TBS with 0.1% Tween-20 (TBST) for one hour at room temperature, followed by incubation with anti-CHD8 primary antibody (rabbit anti-Chd8 N-terminal, Bethyl Laboratories, 1:5000) in 3% non-fat milk powder and 1% BSA in TBST overnight at 4°C. After washing, membrane was incubated with HRP-conjugated secondary antibody (Millipore) for one hour at room temperature. HRP was detected with Clarity ECL reagent (Bio-Rad) and the membrane imaged using a Bio-Rad ChemiDoc system. The membrane was then washed in TBST and incubated overnight at 4°C in 0.05% sodium azide in PBS, before washing and incubation with anti-GAPDH primary antibody (rabbit anti-GAPDH, Abcam, 1:40000) overnight at 4°C. Membrane was probed with HRP-

conjugate and imaged as before. Relative protein quantity was calculated using Bio-Rad ImageLab software.

X-ray Computed tomography

Fixed heads from adult (26 - 27 days old) *Chd8*$^{+/-}$ and *Chd8*$^{+/+}$ mice (n=7 of each from two different litters) were scanned using a GE Locus SP microCT scanner. The specimens were immobilised using cotton gauze and scanned to produce 28µm voxel size volumes, using a X-ray tube voltage of 80kVp and a tube current of 80µA. An aluminium filter (0.05mm) was used to adjust the energy distribution of the X-ray source. Reconstructions of computer tomography scans, images and measurements were done in MicroView 2.5.0 software from Parallax Innovations. Each 3D landmark point was recorded, twice for each sample, using the 3D point recording built-in tool within the same software, with the operator blind to the genotypes. The distances between the landmarks were normalised for each sample to the average of the wild-type littermates. Graphics of the plotted data and statistical analysis were performed using GraphPad Prism version 6.0h for Mac OS X (GraphPad Software, La Jolla California USA, www.graphpad.com). Unpaired student t-tests were applied to analyse the variation between the two groups, for every distance between 2 specific 3D landmark points. Three-dimensional coordinate locations of a total of 22 biological relevant cranial landmarks were chosen based on the landmark list for adult mouse skull proposed by the Richtsmeier lab (http://getahead.psu.edu/landmarks_new.html) (Hill et al., 2009).

3D LANDMARK POINTS LIST:

1.- Nasale: intersection of nasal bones at midline, rostral point

2.- Nasion: intersection of nasal bones at midline, caudal point

3.- Bregma: intersection of frontal bones and parietal bones at midline

4.- Lambda: intersection of parietal bones with anterior aspect of interparietal bone at midline

5.- Intersection of interparietal bones with squamous portion of occipital bone at midline

6.- Anterior-most portion at intersection of premaxillae and nasal bones, left side

7.- Anterior-most portion at intersection of premaxillae and nasal bones, right side

8.- Anterior notch on frontal process lateral to infraorbital fissure, left side

9.- Anterior notch on frontal process lateral to infraorbital fissure, right side

10.- Frontal-squasmosal intersection at temporal crest, left side

11.- Frontal-squasmosal intersection at temporal crest, right side

12.- Joining of squasmosal body to zygomatic process of squasmosal, left side

13.- Joining of squasmosal body to zygomatic process of squasmosal, right side

14.- Intersection of parietal, temporal and interparietal bones, left side

15.- Intersection of parietal, temporal and interparietal bones, right side

16.- Most anterior point of the anterior palatine foramen, left side

17.- Most anterior point of the anterior palatine foramen, right side

18.- Most posterior point of the anterior palatine foramen, left side

19.- Most posterior point of the anterior palatine foramen, right side

20.- Posterior nasal spine, most posterior projection of the posterior nasal spine (palatine bone)

21.- Basion: midsaggital point on the anterior margin of the foramen magnum

22.- Opisthion: midsaggital point on the posterior margin of the foramen magnum

Behavioural assessments

Mice for behavioural testing were housed in standard cages measuring 32×16×14cm with sawdust (Litaspen premium, Datesand Ltd, Manchester), a cardboard shelter and additional bedding material (Sizzlenest, Datesand Ltd, Manchester) with *ad libitum* access to water and

food (Rat and Mouse No. 1 and 3 Maintenance Diet for test and breeding mice respectively, Special Diet Services, Essex, UK). The housing and test rooms were maintained at constant room temperature (21°C) and humidity (45%) and kept under a regular light/dark schedule with lights on from 07:30 to 19:30 hours (light = 270 lux).

Behavioural experiments were conducted between 08:30 and 18:30 in sound-proofed rooms under standard room lighting unless stated otherwise. Behaviours were recorded using a camera positioned above and/or on the side of the test arenas and movement of each mouse tracked using EthoVision (Noldus Information Technologies bv, Wageningen, The Netherlands), or scored manually using Matlab (version 8.5). For automated analyses of the animals' behaviour using Ethovision the following were recorded and analysed: The mean velocity (cm/s) and total distance (cm) travelled in different compartments of the arena, the latency (s) to enter different parts of the arena(s) and duration (s) in each compartment.

After each individual trial of a specific test, boli and urine were removed from the test arena which was cleaned with 1% Anistel® solution (high level surface disinfectant, Trisel Solution Ltd, Cambridgeshire, UK) to remove any odours. Experimenters were blind to the genotype of the animals both during the testing and subsequent scoring of the recorded behaviours.

Different batches of mice were used for (i) recording pup USVs and spontaneous motor behaviours, and (ii) adult behaviours. For the first batch of mice, paw tattoos were administered, immediately after testing on P2, to allow for identification of pups prior to ear notching. All batches of mice were ear notched for permanent identification at P14 and housed in same-sex groups of 2-3 after weaning (P21). For adult behaviours tests were

carried out in the following order: rotarod, grip strength, open field, self-grooming, marble burying, adult social investigation, 3 chamber social approach, light/dark test, olfactory habituation/dishabituation and Morris water maze.

One week before performing the rotarod test, mice were singly-housed to avoid any potential confounds from social and aggressive behaviour hierarchies, which could influence the controlled assessment of social behaviours (Brown, 1953). Sawdust was changed every other week but never on the day before, or the day of, testing and the enrichment (nesting material and house) was changed less regularly to minimize the disruption to the animals. For all social tests, conspecific mice were housed in a separate room to the test mice to ensure the conspecifics were unfamiliar to the test mice. Test mice were never exposed to the same conspecific during testing to ensure novelty.

Ultrasonic vocalisations (USVs) in response to maternal & nest separation: USVs and spontaneous motor behaviours were recorded in pups across 3min sessions in response to social separation from the mother and siblings at P2, 4, 6, 8 and 12, as described previously (39, 40). An ultrasound microphone (Avisoft UltraSoundGate condenser microphone capsule CM16, Avisoft Bioacoustics, Berlin, Germany), sensitive to frequencies of 10–180 kHz, was placed through a hole in the middle of the cover of the sound-attenuating box, approximately 20 cm above the pup which was placed in a glass cup. Vocalisations were recorded using Avisoft Recorder software (Version 3.2). For acoustical analysis, recordings were transferred to Avisoft SASLab Pro (Version 4.40) and a fast fourier transformation (FFT) was conducted. Spectrograms were generated at a frequency resolution of 488 Hz and a time resolution of 1ms. The number of calls for each day of testing were analysed to define the ontogenetic profile of emission in control and mutant pups.

Rotarod: Motor coordination and learning were assessed on a rotating rod (Ugo Basile, Milan, Italy) as described (Wohr et al., 2013), with the exception that the maximum speed was set to 30rpm. The latency to fall for any particular day was calculated as the mean of 2 trials.

Grip strength: To assess the neuromuscular ability of the animals, fore- and hindlimb grip strength was measured using a Linton Grip Strength Meter (MJS Technology Ltd, Stevenage UK). Fore- and hindlimb grip strength was measured 3 times and the mean grip strength of the 3 trials reported (Whittemore et al., 2003).

Open field: Mice were placed facing the wall of a circular open field arena (40 cm diameter) and allowed to freely explore for 10 min. An area of equal distance from the periphery (20 cm diameter), defined as the 'central zone' which is assumed to be the more anxiolytic part of the arena, was virtually drawn in Ethovision.

Self-grooming: This test was carried out as described (McFarlane et al., 2008). The animals were placed in an empty standard housing cage and given 10 minutes to habituate in a dimly lit test room (< 10 lux). The cumulative time spent self-grooming in the following 10 minutes was recorded by an experimenter.

Marble burying: Repetitive digging behaviour to bury marbles was measured, as described by Deacon, 2006 (Deacon, 2006). Repetitive digging behaviour to bury marbles was measured in a dimly lit test room (< 10 lux). Twelve blue glass marbles were arranged in a symmetrical 4×3cm grid on top of 5 cm deep sawdust) in a clean, standard housing cage.

Each mouse was given 30min to freely explore the cage. Eight mice were tested simultaneously and the number of marbles buried at 2.5, 5, 7.5, 10, 20 and 30 min intervals were counted by the experimenter.

Adult social investigation: Social investigation of test mice in response to C57BL/6J sex-matched conspecifics was assessed in adulthood. Test mice were placed into a new standard housing cage containing sawdust and a novel, sex-matched conspecific C57BL/6J mouse introduced to the test cage. The test room was dimly lit (10 lux). Mice were allowed to interact for 3 minutes and the behaviour was recorded. The following behaviours (frequency and duration in s) initiated by the test mouse were scored: anogenital sniffing (direct contact with the anogenital area), body sniffing (sniffing or snout contact with the flank area), head sniffing (sniffing or snout contact with the head/neck/mouth area). No observations of mounting, fighting, tail rattling, and wrestling behaviours were observed. The test room was dimly lit (10 lux).

Light/dark box: a custom-built box of grey acrylic was used with dimensions (44 x 21 x 21 cm). The box was divided into two chambers by a sheet of grey acrylic (21 x 50 cm); a smaller dark chamber (20 lux) that occupied roughly 1/3 of the total box, and a larger lit chamber (80-110 lux) that was lit from above with a bright white light. A small doorway (5 x 7 cm) within the partition allowed the mice to move between chambers freely. Mice were placed in the dark compartment at the start of the 5-min trial. Entry to either compartment was defined as when all four paws of the mouse had entered in one compartment.

Three-chamber social approach: This task was essentially carried out as described (Yang et al., 2011), with 2 alterations: (1) the chamber was not automated and instead Ethovision was

used to track the activity of the test mouse and (2) a novel object (tally counter, Appleton Woods GC101) was placed under the cup in trial 2. The mice were allowed to freely explore the three-chamber apparatus over two 10 min trials. During trial 1, the apparatus was empty and the locomotor activity of the test mouse was tracked using Ethovision. In trial 2, one wired cup containing the inanimate object was placed upside down in one of the side chambers (novel object stimulus) and a novel age and sex-matched conspecific mouse was placed under another wire cup in the other side chamber (novel mouse stimulus), leaving the middle chamber empty. The location of the novel mouse across trials was counterbalanced to minimise any potential confound due to a preference for chamber location.

Olfactory habituation/dishabituation test: This task was essentially carried out as described (Yang and Crawley, 2009). Animals were tested in their home cage, with all the enrichment removed and a fresh cage lid provided just before the trial commenced to minimise the amount of interfering odours (the cage of each mouse was cleaned 3 days prior to testing). Following a 10 min habituation, the mouse was exposed to three odours in turn: water (control/no odor; 50µl), banana essence (non-social; 50µl; Uncle Roy's, Moffat, UK) and urine collected from novel, sex-matched conspecific mice (social, 25µl), each presented on a cotton-tipped wooden applicator 3 times over 2 minutes. Total time (s) spent by the mouse sniffing the cotton buds during each trial was recorded.

General activity measurements

General activity was measured using a running wheel paradigm. Mice were housed individually under a 12h:12h light-dark cycle (lights on at 8am; lights off at 8pm) in a light-, air-, temperature-controlled ventilated cabinet (Arrowmight, Hereford, UK). Running-wheel cages were equipped with an infrared sensor (Bilaney consultant Ltd, Sevenoaks, UK)

connected to a computer. Data were collected in 1-min bins using Clocklab software (Actimetrics, Inc, Wilmette, IL, USA). Mice were continuously monitored undisturbed from the day they were placed in the running wheel cages and their general activity during the light versus dark phase were compared over the first 7 days. Since male and female mice did not show any statistical significant sex difference, data were pooled.

Morris water maze: This task was carried out as described (Grayton et al., 2013) with the exception that the pool was 100 cm in diameter and that 8 hidden and 6 reversal trials were performed. Latency to reach the platform was manually scored for each mouse by an experiment blind to the genotype of the mouse and path length (cm) to reach the platform and speed (cm/s) were extracted from Ethovision. Mean latencies (s) and path lengths (cm) were calculated across the trials within each session for each mouse. To assess the retention of spatial memory, the time spent in the quadrant that had contained the platform (target quadrant) compared to the other quadrants was measured. Conflicting behavioral responses such as floating or thigmotaxis (the amount of time spent swimming in the outer area of the pool defined as a 15 cm wide circular zone adjacent to the wall of the maze) were assessed throughout the trials.

Structural MRI

After completion of adult behavioural tests, mice were terminally anesthetized and intracardially perfused with 30mL of 0.1M PBS containing 10U/mL heparin (Sigma) and 2mM ProHance (a Gadolinium contrast agent) followed by 30mL of 4% paraformaldehyde (PFA) containing 2mM ProHance (Spring et al., 2007). Perfusions were performed at a rate of approximately 60mL/hr. After perfusion, mice were decapitated. The brain and remaining skull structures were incubated in 4% PFA + 2mM ProHance overnight at 4°C then

transferred to 0.1M PBS containing 2mM ProHance and 0.02% sodium azide for at least 7 days prior to MRI scanning. A multi-channel 7.0 Tesla MRI scanner (Varian Inc., Palo Alto, CA) was used to image the brains within skulls. Sixteen custom-built solenoid coils were used to image the brains in parallel (Bock et al., 2005). Parameters used in the anatomical MRI scans: T2- weighted 3D fast spin-echo sequence, with a cylindrical acqusition of k-space, and with a TR of 350 ms, and TEs of 12 ms per echo for 6 echoes, two averages, field-of-view of 20 x 20 x 25 mm^3 and matrix size = 504 x 504 x 630 giving an image with 0.040 mm isotropic voxels (Lerch et al., 2011). The current scan time required for this sequence is ~14 hours. To visualise and compare any changes in the mouse brains the images were linearly (6 parameter followed by a 12 parameter) and non-linearly registered towards a pre-existing atlas (Dorr et al., 2008), and then iteratively linearly and non-linearly aligned to each other to create a population atlas representing the average anatomy of the study sample. The result of the registration is to have all scans deformed into alignment with each other in an unbiased fashion. This allows for the analysis of the deformations needed to take each individual mouse's anatomy into this final atlas space, the goal being to model how the deformation fields relate to genotype (Lerch et al., 2008; Nieman et al., 2006). The jacobian determinants of the deformation fields were then calculated as measures of volume at each voxel. Significant volume changes were then calculated in two ways, 1) on a region basis, and 2) voxelwise. Regional volumes are calculated by warping a pre-existing classified MRI atlas onto the population atlas. This atlas encompasses 159 different structures including, but not limited to, the cortical lobes, large white matter structures (i.e. corpus callosum), ventricles, cerebellum, brain stem, and olfactory bulbs (Dorr et al., 2008; Steadman et al., 2014; Ullmann et al., 2013). Significant differences can then be determined between groups for the 159 different regions in the brain. Voxelwise comparisons were then made between mutants and littermate controls, and multiple comparisons in this study were controlled for

using the False Discovery Rate (Genovese et al., 2002).

Resting-State fMRI

rsfMRI experiments were performed on 15-18 week old mice (n=23 *Chd8*$^{+/+}$; n=19 *Chd8*$^{+/-}$). Animals were prepared for imaging as previously described (Ferrari et al. 2012; Sforazzini et al. 2016). Briefly, mice were anaesthetised using isoflurane (5% induction), intubated and artificially ventilated (2% maintenance). Blood pressure was monitored continuously by cannulating the left femoral artery, also allowing for terminal arterial blood sampling. Administration of isoflurane was ceased after surgery and substituted with halothane (0.75%). 45 mins after isoflurane cessation functional data acquisition commenced. Throughout each imaging session mean arterial blood pressure was recorded continuously.

In vivo images were obtained using a 7.0 T MRI scanner (Bruker Biospin, Milan), as previously described (Liska et al. 2016). Signal transmission and reception were achieved using a 72mm birdcage transmit coil and a 4-channel solenoid coil. For each session, high resolution anatomical images were acquired using a fast spin echo sequence based on the following parameters: repetition time (TR)/echo time (TE) 5500/60ms, matrix 192 x 192, field of view 2 x 2cm^3, 24 coronal slices, and slice thickness 0.5mm. Co-centred BOLD rsfMRI time series were acquired using an echo planar imaging (EPI) sequence with the following parameters: TR/TE 1200/15ms, flip angle 30°, matrix 100 x 100, field of views 2 x 2cm^2, 24 coronal slices, slice thickness 0.5mm, 500 volumes and 10min total acquisition time. MRI raw data, templates and code employed to generate functional maps are available to readers by contacting AG.

Functional Connectivity Analyses

To allow for T_1 equilibration effects, the first 20 volumes of rsfMRI data were removed. The time series were then despiked, corrected for motion and spatially normalised to an in-house mouse brain template (Sforazzini et al., 2014). Normalised data had a spatial resolution of 0.1042 x 0.1042 x 0.5mm^3 (192 x 192 x 24 matrix). Mean ventricular signal (averaged rsfMRI time course within a reference ventricular mask) and head motion traces were regressed out of each time series. No genotype-dependent differences were observed in ventricular volume, as measured by the dimensions of individual ventricular masks. All rsfMRI time series were then spatially smoothed (full width at half maximum of 0.6mm) and band-pass filtered using a frequency window of 0.01-0.1Hz.

To identify brain regions displaying genotype-dependent differences in functional connectivity in an unbiased manner, we calculated global rsfMRI connectivity maps for all subjects, as described previously in detail (Liska et al., 2017; Liska et al., 2015). A previously described seed-based approach was then used to examine between-group differences in the intensity and scope of long-range rsfMRI correlation networks (Sforazzini et al., 2016).

Tissue Collection and Processing

Pups were weighed and then sacrificed, while embryos were collected by dissection in ice-cold PBS, excess PBS drained and whole embryos weighed. Brains were then dissected from the skull in ice-cold PBS and cut below the brain stem. Brains were then immediately drained on paper towels using a slotted spoon and wet weights determined using a fine scale. Brain weights were normalised to body weight and group differences were calculated using unpaired students t-test. Brains were returned to PBS and wholemount pictures taken on a Nikon SMZ1500 stereo-microscope equipped with a Nikon DS-Fi1 camera head, before post-fixation in 4% PFA at 4°C for 24h. After fixing, brains were dehydrated and paraffin

embedded. Paraffin blocks were then cut into 10µm thick coronal sections and mounted such that each slide contained three adjacent sections.

Immunohistochemistry

Coronal brain sections were re-hydrated using standard protocols. Antigen retrieval was conducted by heating slides in 10mM Sodium Citrate solution (pH6) for 20mins and cooled on ice. Endogenous peroxidases were blocked by incubating in 3% H_2O_2 and 10% MeOH in PBS for 15mins. Sections were then washed in 0.2% Triton X-100 (Sigma-Aldrich) in PBS (PBT2) for 5 mins and blocked using 10% heat-inactivated normal goat serum (GS) and 2% gelatin in PBT2 for 1 hour. Sections were incubated in 5% GS in PBT2 containing primary antibody (rabbit anti-phosphohistone 3B (Cell Signaling, 1/100)). overnight at 4°C. After incubation with primary antibody, sections were incubated in biotinylated anti-rabbit immunoglobulin secondary antibody (Dako, 1/200) in 5% GS in PBT2. Samples were washed in PBS and incubated with Avidin/biotin complex (ABC, Vector) in PBS for 1 hour. Sections were developed using 0.025% DAB and 0.03% H_2O_2 in PBS for 10mins before washing in running water and counterstaining using Ehrlich's Hemotoxylin solution. Slides were mounted onto coverslips using DPX (Sigma-Aldrich). Images were acquired on a Nikon 80i microscope equipped with a Nikon 5M pixel Nikon DS digital cameras. Images were processed using Adobe Photoshop and Illustrator.

RNA Extraction and qRT-PCR analysis.

To extract RNA, dissected cortices were lysed in 600µl Trizol (Life Technologies). RNA was then purified and DNase treated using the Direct-zol RNA MiniPrep kit (Zymo Research) according to the manufacturer's instructions. For qRT-PCR, cDNA was synthesized using 50ng RNA from 4 biological replicates per condition with the Precision nanoScript 2 Reverse

Transcription Kit (PrimerDesign Ltd.) according to the manufacturer's recommendations. qRT-PCRs were performed on a Stratagene Mx3000p (Agilent Technologies) using PrecisionPlus-MX 2x qPCR Mastermix with SYBR green (PrimerDesign Ltd.) and primers against *Chd8* exon 3-4 (FW: CAG AGG AGG AGG GTG AAA AGA AAC, Rev: GAG TTG TCA GAC GAT GTG TTA CGC) or *Chd8* exon 1-2 (FW: TGA AGC CTG CAG TTA CAC TGA CGT, Rev: CTG CGG CTG TGG CTG TGG TT). *Canx* and *Sdha* (E12.5) and *Gapdh* and *Eifa* (P5) were used as endogenous control genes as determined by prior geNorm (Primerdesign, UK) analysis for the respective sample sets, and relative expression levels were calculated using the $2^{-\Delta\Delta C_T}$ method.

RNA Sequencing

For RNA-sequencing mRNA was isolated from micro-dissected cortices at E12.5 (both hemispheres) and P5 (one hemisphere) and reverse transcribed into cDNA (n=3 per experimental group). cDNA was end-repaired, adaptor-ligated and A-tailed. Paired-end sequencing was performed on the Illumina HiSeq 4000 platform. Quality of the raw sequencing data was checked using FastQC version 0.11.2 (Andrews 2010; available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and trimming of adaptor sequences was performed using Trim Galore! version 0.4.1 (Krueger 2012; available at: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Reads were aligned to the mouse genome (GRCm38.p4) using Tophat version 2.1.0 and aligned reads were counted using FeatureCounts version 1.5.0 (Kim et al., 2013; Liao et al., 2014). Differential expression testing was performed using DESeq2 version 1.10.1, as previously described (Love et al., 2014). Gene ontology analysis and functional classification was performed using DAVID with all detected DEGs below a 0.05 FDR (Huang da et al., 2009). Putative regulatory transcription factors were determined using Enrichr using the "ENCODE and

ChEA Consensus TFs from ChIP-X" database with all DEGs below a 0.05 FDR (Chen et al., 2013). The R package ggplot2 version 2.1.0 was used to generate volcano plots and DESeq2 normalised read count plots for individual genes. The list of ASD associated genes used for overlap with P5 DEGs was obtained from the SFARI Gene Scoring module (https://gene.sfari.org/autdb/HG_Home.do). RNA-seq data have been deposited into GEO, accession number GSE81103.

Statistical Analysis

Data are reported as Mean±SEM and graphs show all individual data points where feasible. Significant p-values and statistical tests are reported in the results section. Figure legends provide full details of all relevant statistical parameters including group sizes. Statistical analyses were performed either with SPSS (Version 22, IBM, Armonk, USA) or GraphPad Prism (Version 6, GraphPad Software, La Jolla, California, USA). All analyses were performed blind to genotype.

Behaviour: Data were analysed using either a between-subjects ANOVA or a 2-way repeated measures ANOVA, as appropriate. If there was no statistically significant sex difference, data were pooled. When the appropriate ANOVA showed a significant effect for a particular task, student's t-tests were used as post-hoc analyses, as there were only 2 groups for comparison. Cohort details can be found in the methods, group sizes are stated in the figure legend.

Proliferation: Phosphohistone 3B-positive cells lining the ventricular surface of the dorsal cortex were counted and normalised to the length of ventricular surface. These were quantified on both sides of the brain in three consecutive sections and averaged to calculate the number of phosphohistone 3B-positive cells per μm of ventricular surface in the dorsal

cortex. Group differences were calculated using unpaired student's t-test.

μCT analysis: Each 3D landmark point was recorded twice for each sample and distances between landmark points normalised to the average of the wildtype controls. Group differences for distances between two specific 3D landmark points were calculated using unpaired student t-tests.

MRI analyses: Processing of raw data is described in detail in the relevant method sections. For structural MRI, significant differences were determined between groups for the 159 different regions in the brain. Voxelwise comparisons were made between mutants and littermate controls, group differences were calculated using unpaired student's t-test and multiple comparisons were controlled for using a False Discovery Rate (FDR< 0.15). Exact p-values can be found in Extended Data Table 2-1.

For rsfMRI studies, group-level differences in connectivity distributions were calculated using 2-tailed student's t-test ($p<0.05$, family-wise error cluster-corrected, with cluster-defining threshold of $t_{24}>2.06$, $p<0.05$) and multiple comparisons were controlled for using an FDR<0.05.

RNAseq: Processing of raw data and differential expression testing is described in the methods section. Multiple comparisons were controlled for using an FDR<0.05. Exact p-values and FDR adjusted p-values for all differentially expressed genes are listed in Tables 5-1 and 5-2.

ACKNOWLEDGEMENTS

This work was supported by research grants from the Medical Research Council (MR/K022377/1, MAB and CF), Simons Foundation (SFARI #344763, MAB; SFARI #400101, AG), Ontario Brain Institute's POND programme (JPL), and BBSRC (BB/K008668/1, PFW). SH and RE were supported by the King's Bioscience Institute and the Guy's and St Thomas' Charity Prize PhD Programme in Biomedical and Translational Science. AG was supported by a 2017 NARSAD Independent Investigator Grant from the Brain and Behavior Research Foundation. We thank John Whittingham, Alex Donovan and BSU staff for technical assistance and Chris Healy for μCT scans. We acknowledge the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z) for the generation of the RNA sequencing data and Drs. Brian Nieman and Leigh Spencer Noakes for their MRI sequences.

REFERENCES

APA (2013). Diagnostic and statistical manual of mental disorders, 5th edn (Arlington, VA: American Psychiatric Publishing).

Bagri, A., Marin, O., Plump, A.S., Mak, J., Pleasure, S.J., Rubenstein, J.L., and Tessier-Lavigne, M. (2002). Slit proteins prevent midline crossing and determine the dorsoventral position of major axonal pathways in the mammalian forebrain. Neuron *33*, 233-248.

Barnard, R.A., Pomaville, M.B., and O'Roak, B.J. (2015). Mutations and Modeling of the Chromatin Remodeler CHD8 Define an Emerging Autism Etiology. Front Neurosci *9*, 477.

Bedford, R., Pickles, A., and Lord, C. (2016). Early gross motor skills predict the subsequent development of language in children with autism spectrum disorder. Autism research : official journal of the International Society for Autism Research *9*, 993-1001.

Belmonte, M.K., Allen, G., Beckel-Mitchener, A., Boulanger, L.M., Carper, R.A., and Webb, S.J. (2004). Autism and abnormal development of brain connectivity. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 9228-9231.

Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A.T.*, et al.* (2014). Disruptive CHD8 mutations define a subtype of autism early in development. Cell *158*, 263-276.

Blatt, G.J. (2012). The neuropathology of autism. Scientifica (Cairo) *2012*, 703675.

Bock, N.A., Nieman, B.J., Bishop, J.B., and Mark Henkelman, R. (2005). In vivo multiple-mouse MRI at 7 Tesla. Magn Reson Med *54*, 1311-1316.

Bourgeron, T. (2015). From the genetic architecture to synaptic plasticity in autism spectrum disorder. Nat Rev Neurosci *16*, 551-563.

Brown, R.Z. (1953). Social behaviour, reproduction and population changes in the house mouse. Ecol Monogr *23*, 217-240.

Bruneau, N., Dourneau, M.C., Garreau, B., Pourcelot, L., and Lelord, G. (1992). Blood flow response to auditory stimulations in normal, mentally retarded, and autistic children: a preliminary transcranial Doppler ultrasonographic study of the middle cerebral arteries. Biological psychiatry *32*, 691-699.

Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Molecular cell *15*, 57-67.

Cerliani, L., Mennes, M., Thomas, R.M., Di Martino, A., Thioux, M., and Keysers, C. (2015). Increased Functional Connectivity Between Subcortical and Cortical Resting-State Networks in Autism Spectrum Disorder. JAMA psychiatry *72*, 767-777.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC bioinformatics *14*, 128.

Cherkassky, V.L., Kana, R.K., Keller, T.A., and Just, M.A. (2006). Functional connectivity in a baseline resting-state network in autism. Neuroreport *17*, 1687-1690.

Chinello, A., Di Gangi, V., and Valenza, E. (2016). Persistent primary reflexes affect motor acts: Potential implications for autism spectrum disorder. Research in developmental disabilities.

Constantino, J.N., and Charman, T. (2016). Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. Lancet Neurol *15*, 279-291.

Cotney, J., Muhle, R.A., Sanders, S.J., Liu, L., Willsey, A.J., Niu, W., Liu, W., Klei, L., Lei, J., Yin, J.*, et al.* (2015). The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. Nature communications *6*, 6404.

Croen, L.A., Zerbo, O., Qian, Y., Massolo, M.L., Rich, S., Sidney, S., and Kripke, C. (2015). The health status of adults on the autism spectrum. Autism : the international journal of research and practice *19*, 814-823.

Deacon, R.M. (2006). Digging and marble burying in mice: simple methods for in vivo identification of biological impacts. Nature protocols *1*, 122-124.

Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M.*, et al.* (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular psychiatry *19*, 659-667.

Donovan, A.P., and Basson, M.A. (2017). The neuroanatomy of autism - a developmental perspective. J Anat *230*, 4-15.

Dorr, A.E., Lerch, J.P., Spring, S., Kabani, N., and Henkelman, R.M. (2008). High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice. NeuroImage *42*, 60-69.

Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A., and Tsai, L.H. (2016). Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat Neurosci.

Ecker, C. (2016). The neuroanatomy of autism spectrum disorder: An overview of structural neuroimaging findings and their translatability to the clinical setting. Autism.

Edgar, J.C., Fisk Iv, C.L., Berman, J.I., Chudnovskaya, D., Liu, S., Pandey, J., Herrington, J.D., Port, R.G., Schultz, R.T., and Roberts, T.P. (2015). Auditory encoding abnormalities in children with autism spectrum disorder suggest delayed development of auditory cortex. Molecular autism *6*, 69.

Genovese, C.R., Lazar, N.A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage *15*, 870-878.

Gompers, A.L., Su-Feher, L., Ellegood, J., Copping, N.A., Riyadh, M.A., Stradleigh, T.W., Pride, M.C., Schaffler, M.D., Wade, A.A., Catta-Preta, R.*, et al.* (2017). Germline Chd8 haploinsufficiency alters brain development in mouse. Nat Neurosci *advance online publication*.

Grayton, H.M., Missler, M., Collier, D.A., and Fernandes, C. (2013). Altered social behaviours in neurexin 1alpha knockout mice resemble core symptoms in neurodevelopmental disorders. PLoS One *8*, e67114.

Hill, C.A., Sussan, T.E., Reeves, R.H., and Richtsmeier, J.T. (2009). Complex contributions of Ets2 to craniofacial and thymus phenotypes of trisomic "Down syndrome" mice. Am J Med Genet A *149A*, 2158-2165.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E.*, et al.* (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216-221.

Jeste, S.S., and Tuchman, R. (2015). Autism Spectrum Disorder and Epilepsy: Two Sides of the Same Coin? Journal of child neurology *30*, 1963-1971.

Just, M.A., Cherkassky, V.L., Keller, T.A., and Minshew, N.J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. Brain : a journal of neurology *127*, 1811-1821.

Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., Suyama, M., Takumi, T., Miyakawa, T., and Nakayama, K.I. (2016). CHD8 haploinsufficiency results in autistic-like phenotypes in mice. Nature *537*, 675-679.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Lerch, J.P., Carroll, J.B., Spring, S., Bertram, L.N., Schwab, C., Hayden, M.R., and Henkelman, R.M. (2008). Automated deformation analysis in the YAC128 Huntington disease mouse model. NeuroImage *39*, 32-39.

Lerch, J.P., Sled, J.G., and Henkelman, R.M. (2011). MRI phenotyping of genetically altered mice. Methods Mol Biol *711*, 349-361.

Lewandoski, M., and Martin, G.R. (1997). Cre-mediated chromosome loss in mice. Nat Genet *17*, 223-225.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Liska, A., Bertero, A., Gomolka, R., Sabbioni, M., Galbusera, A., Barsotti, N., Panzeri, S., Scattoni, M.L., Pasqualetti, M., and Gozzi, A. (2017). Homozygous Loss of Autism-Risk Gene CNTNAP2 Results in Reduced Local and Long-Range Prefrontal Functional Connectivity. Cereb Cortex, 1-13.

Liska, A., Galbusera, A., Schwarz, A.J., and Gozzi, A. (2015). Functional connectivity hubs of the mouse brain. NeuroImage *115*, 281-291.

Liska, A., and Gozzi, A. (2016). Can Mouse Imaging Studies Bring Order to Autism Connectivity Chaos? Front Neurosci *10*, 484.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Maness, P.F., and Schachner, M. (2007). Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. Nature neuroscience *10*, 19-26.

Marco, E.J., Hinkley, L.B., Hill, S.S., and Nagarajan, S.S. (2011). Sensory processing in autism: a review of neurophysiologic findings. Pediatr Res *69*, 48R-54R.

McFarlane, H.G., Kusek, G.K., Yang, M., Phoenix, J.L., Bolivar, V.J., and Crawley, J.N. (2008). Autism-like behavioral phenotypes in BTBR T+tf/J mice. Genes, brain, and behavior *7*, 152-163.

Merner, N., Forgeot d'Arc, B., Bell, S.C., Maussion, G., Peng, H., Gauthier, J., Crapper, L., Hamdan, F.F., Michaud, J.L., Mottron, L.*, et al.* (2016). A de novo frameshift mutation in chromodomain helicase DNA-binding domain 8 (CHD8): A case report and literature review. Am J Med Genet A.

Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V.*, et al.* (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature *485*, 242-245.

Nieman, B.J., Flenniken, A.M., Adamson, S.L., Henkelman, R.M., and Sled, J.G. (2006). Anatomical phenotyping in the brain and skull of a mutant mouse by magnetic resonance imaging and computed tomography. Physiol Genomics *24*, 154-162.

O'Roak, B.J., Stessman, H.A., Boyle, E.A., Witherspoon, K.T., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J.B., Nickerson, D.A.*, et al.* (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. Nature communications *5*, 5595.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K.*, et al.* (2012a). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science *338*, 1619-1622.

O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D.*, et al.* (2012b). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature *485*, 246-250.

Picci, G., Gotts, S.J., and Scherf, K.S. (2016). A theoretical rut: revisiting and critically evaluating the generalized under/over-connectivity hypothesis of autism. Developmental science *19*, 524-549.

Platt, R.J., Zhou, Y., Slaymaker, I.M., Shetty, A.S., Weisbach, N.R., Kim, J.A., Sharma, J., Desai, M., Sood, S., Kempton, H.R.*, et al.* (2017). Chd8 Mutation Leads to Autistic-like Behaviors and Impaired Striatal Circuits. Cell Rep *19*, 335-350.

Scott-Van Zeeland, A.A., Abrahams, B.S., Alvarez-Retuerto, A.I., Sonnenblick, L.I., Rudie, J.D., Ghahremani, D., Mumford, J.A., Poldrack, R.A., Dapretto, M., Geschwind, D.H.*, et al.* (2010). Altered functional connectivity in frontal lobe circuits is associated with variation in the autism risk gene CNTNAP2. Sci Transl Med *2*, 56ra80.

Sforazzini, F., Bertero, A., Dodero, L., David, G., Galbusera, A., Scattoni, M.L., Pasqualetti, M., and Gozzi, A. (2016). Altered functional connectivity networks in acallosal and socially impaired BTBR mice. Brain structure & function *221*, 941-954.

Sforazzini, F., Schwarz, A.J., Galbusera, A., Bifone, A., and Gozzi, A. (2014). Distributed BOLD and CBV-weighted resting-state networks in the mouse brain. NeuroImage *87*, 403-415.

Sinclair, D., Oranje, B., Razak, K.A., Siegel, S.J., and Schmid, S. (2017). Sensory processing in autism spectrum disorders and Fragile X syndrome-From the clinic to animal models. Neuroscience and biobehavioral reviews *76*, 235-253.

Spring, S., Lerch, J.P., and Henkelman, R.M. (2007). Sexual dimorphism revealed in the structure of the mouse brain using three-dimensional magnetic resonance imaging. NeuroImage *35*, 1424-1433.

Steadman, P.E., Ellegood, J., Szulc, K.U., Turnbull, D.H., Joyner, A.L., Henkelman, R.M., and Lerch, J.P. (2014). Genetic effects on cerebellar structure across mouse models of autism

using a magnetic resonance imaging atlas. Autism research : official journal of the International Society for Autism Research *7*, 124-137.

Stessman, H.A., Xiong, B., Coe, B.P., Wang, T., Hoekzema, K., Fenckova, M., Kvarnung, M., Gerdts, J., Trinh, S., Cosemans, N.*, et al.* (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nature genetics *49*, 515-526.

Stolerman, E.S., Smith, B., Chaubey, A., and Jones, J.R. (2016). CHD8 intragenic deletion associated with autism spectrum disorder. Eur J Med Genet *59*, 189-194.

Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J.*, et al.* (2014). CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. Proc Natl Acad Sci U S A *111*, E4468-4477.

Talkowski, M.E., Rosenfeld, J.A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A.M.*, et al.* (2012). Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell *149*, 525-537.

Tavassoli, T., Miller, L.J., Schoen, S.A., Nielsen, D.M., and Baron-Cohen, S. (2014). Sensory over-responsivity in adults with autism spectrum conditions. Autism : the international journal of research and practice *18*, 428-432.

Thompson, B.A., Tremblay, V., Lin, G., and Bochar, D.A. (2008). CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. Mol Cell Biol *28*, 3894-3904.

Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A., and Warman, M.L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). Biotechniques *29*, 52, 54.

Ullmann, J.F., Watson, C., Janke, A.L., Kurniawan, N.D., and Reutens, D.C. (2013). A segmentation protocol and MRI atlas of the C57BL/6J mouse neocortex. NeuroImage *78*, 196-203.

Whittemore, L.A., Song, K., Li, X., Aghajanian, J., Davies, M., Girgenrath, S., Hill, J.J., Jalenak, M., Kelley, P., Knight, A.*, et al.* (2003). Inhibition of myostatin in adult mice increases skeletal muscle mass and strength. Biochem Biophys Res Commun *300*, 965-971.

Wohr, M., Silverman, J.L., Scattoni, M.L., Turner, S.M., Harris, M.J., Saxena, R., and Crawley, J.N. (2013). Developmental delays and reduced pup ultrasonic vocalizations but normal sociability in mice lacking the postsynaptic cell adhesion protein neuroligin2. Behavioural brain research *251*, 50-64.

Yang, M., and Crawley, J.N. (2009). Simple behavioral assessment of mouse olfaction. Current protocols in neuroscience / editorial board, Jacqueline N Crawley [et al] *Chapter 8*, Unit 8 24.

Yang, M., Silverman, J.L., and Crawley, J.N. (2011). Automated three-chambered social approach task for mice. Current protocols in neuroscience / editorial board, Jacqueline N Crawley [et al] *Chapter 8*, Unit 8 26.

Yuan, C.C., Zhao, X., Florens, L., Swanson, S.K., Washburn, M.P., and Hernandez, N. (2007). CHD8 associates with human Staf and contributes to efficient U6 RNA polymerase III transcription. Mol Cell Biol *27*, 8729-8738.

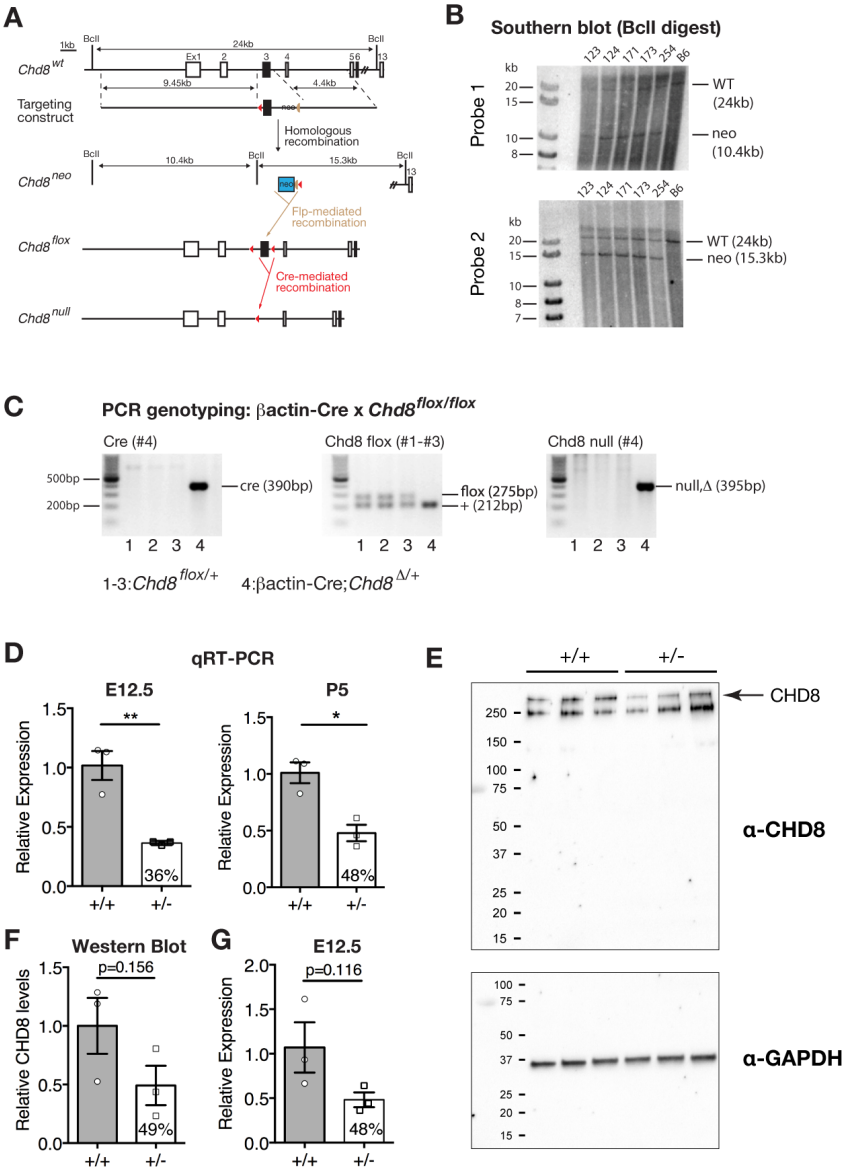


Figure 1. Construction and validation of the *Chd8* conditional and null alleles.
A) Diagrammatic representations of the wildtype (wt) *Chd8* locus (*Chd8*^{wt}), targeting construct, targeted (neo), conditional (flox) and null alleles. Approximate genomic distances are indicated in kilobases (kb), exons are denoted by boxes labelled Ex1 to 13, and Southern blot probes (P1,P2) and PCR primers (#1- #4) are indicated. The 5' long homology arm is shown in red and the 3' short homology arm in green. The neomycin resistance cassette (neo) is shown as a blue box, the floxed exon 3 by a black box, loxP sites by red triangles and frt sites by tan semi-ovals. BclI restriction enzyme sites are labelled.
B) Southern blot of genomic DNA digested with BclI from embryonic stem (ES) cell clones (123, 124 etc.) and wildtype C57BL/6J (B6) cells, hybridised with P1 are shown, with molecular weight markers in the left hand lane. The wildtype allele (WT) gives a 24kb band, whilst the targeted allele gives a band of approximately 10.4kb. Southern blot of genomic DNA digested with BclI from embryonic stem (ES) cell clones as indicated and wildtype B6 cells, hybridised with P2 are shown, with molecular weight markers in the left hand lane. The

wildtype allele (WT) gives a 24kb band, whilst the targeted allele gives a band of approximately 15.3kb.

C) PCR genotyping of genomic DNA extracted from mouse pups from a cross between a heterozygous general deleter *βactin-Cre* transgenic mouse and a *Chd8*$^{flox/flox}$ mouse. Results from PCR reactions to detect the Cre transgene, distinguish the *Chd8*^{flox} and wildtype alleles from each other, and amplify the null allele are shown. Note the loss of the flox allele, with the gain of the null allele in the Cre^{+} pup (lane 4).

D) Quantitative RT-PCR for *Chd8* on mRNA extracted from *Chd8* heterozygous mouse neocortices at E12.5 and P5 and littermate controls using primers spanning the exon3-4 boundary. *Chd8* expression levels in heterozygous mice are significantly reduced to 36% of wildtype controls at E12.5 (**p=0.0059, t=5.346, df=4, student's t-test, n=3 per genotype) and 48% at P5 (*p=0.0105, t=4.538, df=4, student's t-test, n=3 per genotype).

E) Western blot on lysates from *Chd8* heterozygous E12.5 neocortices and littermate controls. Upper panel: The band for full-length CHD8 (arrow, ~290kDa) was quantified in F. Note the absence of detectable levels of truncated protein products for CHD8. Lower panel: Western blot for the loading control GAPDH.

F) Quantification of CHD8 protein levels normalised to GAPDH as shown in E. CHD8 protein levels in heterozygous mice is 49% compared to wildtype littermates (p=0.156, student's t-test, n=3 per genotype).

G) Quantitative RT-PCR for *Chd8* on mRNA extracted from *Chd8* heterozygous mouse neocortices at E12.5 and littermate controls using primers spanning the exon1-2 boundary (p=0.116, student's t-test, n=3 per genotype). Samples are the same as in D.

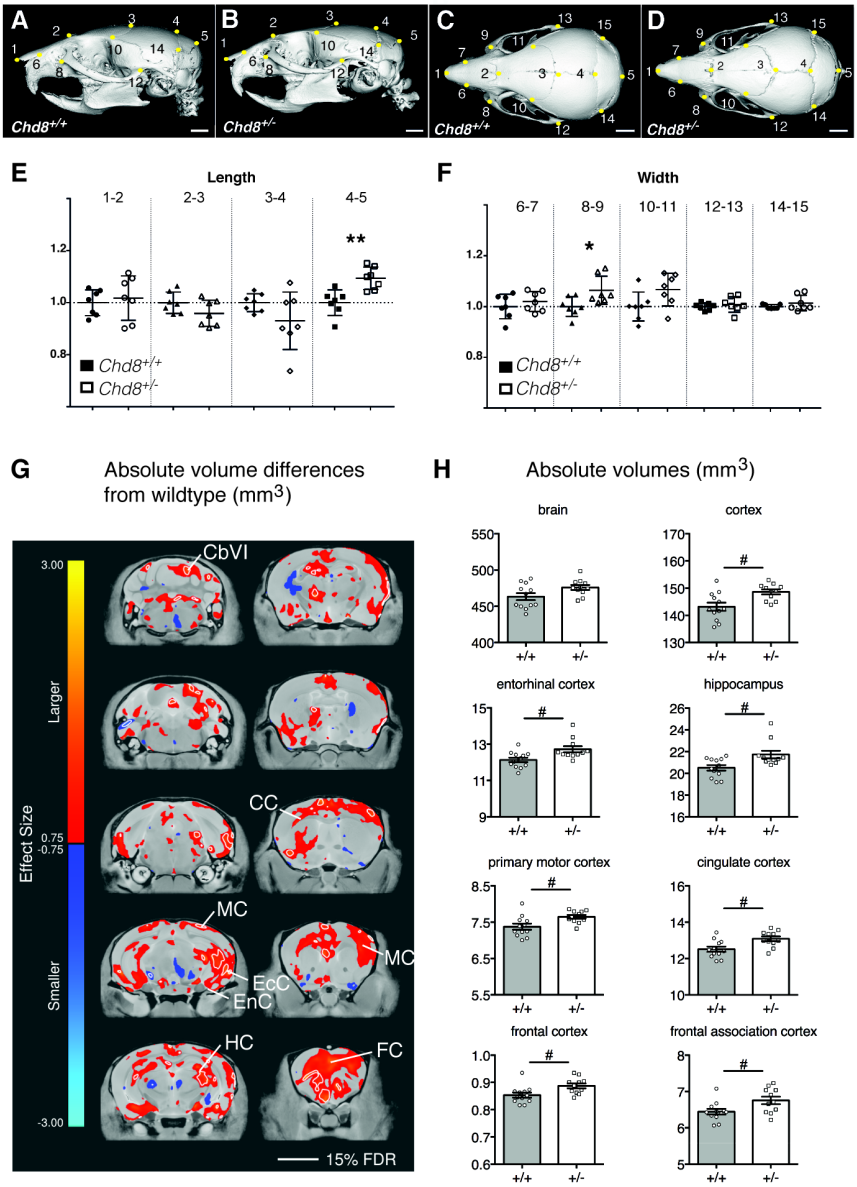


Figure 2. Hypertelorism and mild megalencephaly in *Chd8*$^{+/-}$ mice.

A-D) Representative lateral (A,B) and dorsal (C,D) µCT views of 3D reconstructed skulls from mice with the indicated genotypes. Landmarks from 1 to 15 are indicated by yellow dots. Scale bars = 2mm.

E,F) Graphs for measurements between indicated landmarks, normalised to average measurements from corresponding wildtype littermates. Mean±SEM; landmarks 4-5: p=0.0025, t=3.797; landmarks 8-9: p=0.0273, t=2.512; df=12, student's t-test, n=7 per genotype.

G) High-resolution 7T structural MRI coronal images of *Chd8*$^{+/-}$ brains from posterior (top left) to anterior (bottom right) are shown.

Absolute volumetric differences in size, relative to wildtype controls are coloured according to the scale on the left. Effect size is measured in units of standard deviation. Some regions with enlarged volumes are labelled as follows: CbVI – cerebellar lobule VI, MC – motor

cortex, EcC – ectorhinal cortex, EnC – entorhinal cortex, HC – hippocampus, CC - Cingulate cortex, FC – frontal association cortex.

H) Absolute volumes (mm^3) are plotted for whole brain, neocortex and several other brain regions for the different genotypes as indicated. #FDR<0.15, student's t-tests: brain: p=0.0484, t=2.096; cortex: p=0.0055, t=3.093; entorhinal cortex: p=0.011, t=2.788; hippocampus: p=0.0091, t=2.873 primary motor cortex: p=0.0126, t=2.727; cingulate cortex: p=0.0074, t=2.965; frontal cortex: p=0.0154, t=2.639; frontal association cortex: p=0.0238, t=2.438; df=21, *Chd8*$^{+/-}$: n=11, *Chd8*$^{+/+}$: n=12.

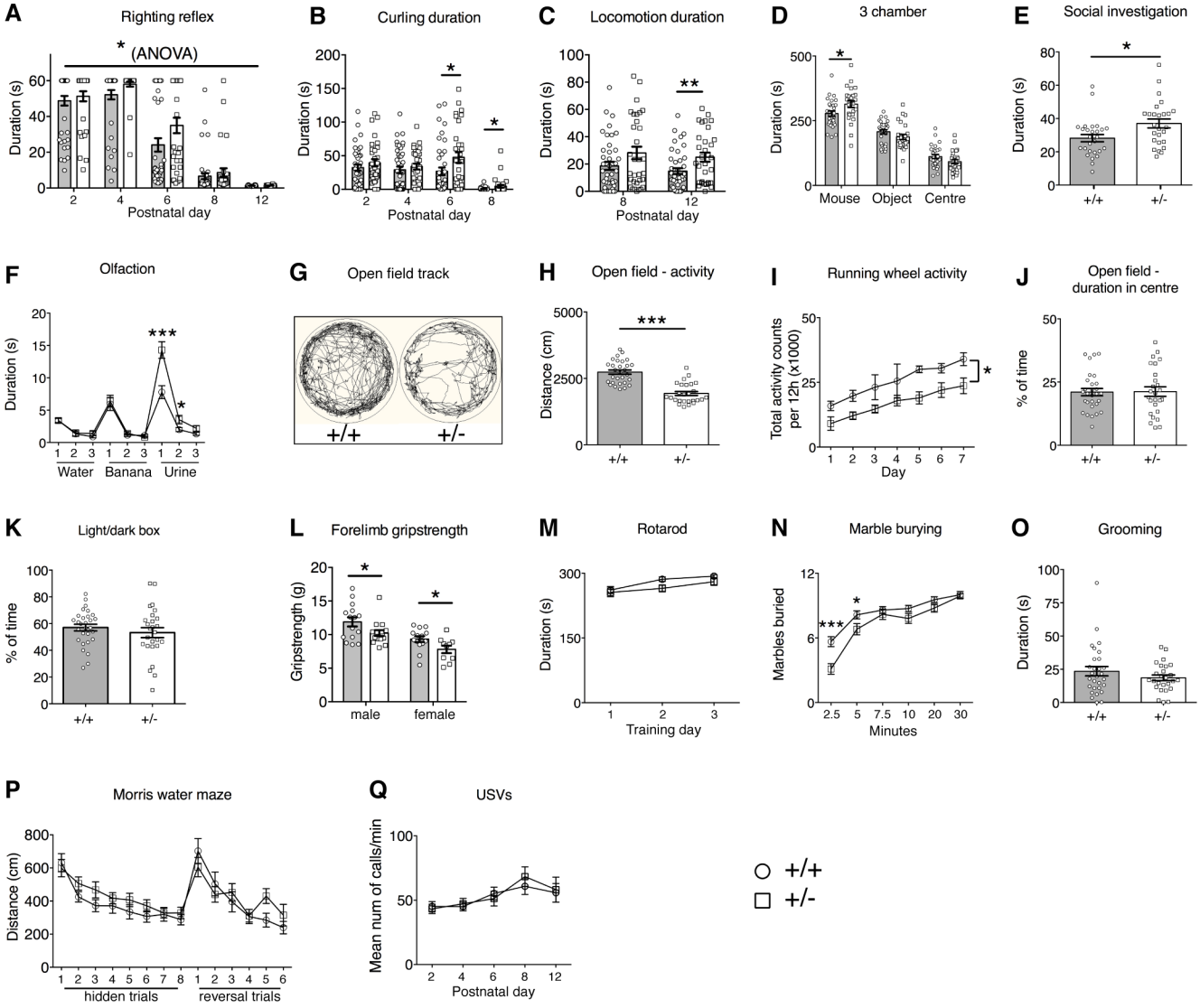


Figure 3. Complex behavioural abnormalities in *Chd8* heterozygous mice.
A-Q) Behavioural assessments of a cohort of adult *Chd8*$^{+/-}$ (+/-, n=26) and *Chd8*$^{+/+}$ (+/+, n=29) and of pup *Chd8*$^{+/-}$ (+/-, n=32) and *Chd8*$^{+/+}$ (+/+, n=42) animals.

A) The development of the righting reflex in pups at the indicated postnatal days. Pups failing to right by the end of the 60s test period were given a score of 60s. Note the significant delay in the acquisition of the full righting reflex response in *Chd8*$^{+/-}$ animals compared to littermate controls. Mean ±1SEM; *p=0.014 (one-way repeated-measures ANOVA: f(1,72)=6.36 (between-subjects effect)).
B) The duration, in seconds, pups spend rolling on their back (curling) as recorded during the analysis of spontaneous movements during USV recordings. Note that *Chd8*$^{+/-}$ mice spent significantly more time curling at P6 and P8 compared to littermate controls. Mean ± SEM; P6 *p=0.0312, P8 *p=0.0354 (one-way repeated-measures ANOVA: f(1,72)=12.64 p=0.001 (between-subjects effect), with student's t-test as post-hoc analysis (p6: df=72, t=2.197, P8: df=72, t=2.145).
C) The duration, in seconds, pups spend in locomotion as recorded during the analysis of spontaneous movements during USV recordings. At P12 *Chd8*$^{+/-}$ animals spent significantly more time in locomotion as compared to littermate controls. Mean ±1SEM, **p=0.009 (one-

way repeated measures ANOVA: ($f(1,72)=7.33$ $p=0.008$ (between-subjects effect), with student's t-test as post-hoc analysis df=72, t=2.687).

D) The duration, in seconds, spent in each chamber of the three-chamber sociability test. All mice spent a significantly higher proportion of time in the chamber with the age- and sex-matched stranger con-specific mouse compared to the other chambers. Mean±SEM; *p=0.029 (between-subjects ANOVA: $f(1,53)=5.031$).

E) Duration, in seconds, of social investigation over a three-minute period. Social investigation was defined as the combined total duration of head, body and anogenital sniffing of a conspecific mouse. Mean ±SEM; *p=0.015 (between-subjects ANOVA $f(1,52)=6.307$).

F) Graph demonstrating the performance in the olfactory habituation/dishabituation test. Mean±SEM;*p=0.03,**p=0.0002 (repeated-measures ANOVA: $f(2.85,145.23)=9.24$ p=0.00002, with student's t-test as post-hoc analysis **df=53,t=4.04, *df=53,t=2.23).

G) Representative ethovision tracks of a *Chd8*$^{+/-}$ (+/-) and *Chd8*$^{+/+}$ (+/+) animal plotting their movements during the 10-minute open field task.

H) The total distance travelled in the outer part of the open field arena over a 10-minute time-period. Mean±SEM; ***p=$2x10^{-9}$ (between-subjects ANOVA: $f(1,53)=52.72$).

I) The total activity counts per 12h period on running wheels in the homecage during 7 days of dark-phase recording. Mean±SEM; *p=0.019 (repeated-measures ANOVA: $f(1,7)=9.12$, between-subjects effect).

J) The percentage of time spent in the centre of the open field arena during the 10-minute test. Mean±SEM (between-subjects ANOVA: $f(1,53)=0.007$, p=0.93).

K) The percentage of time spent in the light chamber during the 5minute light/dark test. Mean±SEM (between-subjects ANOVA: $f(1,51)=0.824$, p=0.368).

L) The average of 3 measurements of forelimb grip strength on a Linton Grip Strength meter. Mean ± SEM, males: *p=0.045 (between-subjects ANOVA: $f(1,29)=4.371$) females: *p=0.042 (between-subjects ANOVA: $f(1,22)=4.677$).

M) The mean latency of mice to fall from the rotarod. Mean±SEM (repeated-measures ANOVA: $f(1.644,102)=0.620$, p=0.540).

N) The average number of marbles buried, out of a maximum of 12, within a 30-minute time period. Mean±SEM; *p=0.04, ***p=0.0004, (repeated-measures ANOVA: $f(3.66,265)=4.70$ p=0.002, with student's t-test as post-hoc analysis * df=53,t=2.12, ***df=53, t=3.79).

O) The duration, in seconds, mice spent self-grooming during the 10-minute self-grooming test. Mean±SEM (between-subjects ANOVA: $f(1,51)=1.21$, p=0.28).

P) Graph plotting the average distance swum for 4 trials daily over 8 consecutive training days to find the hidden platform (hidden trials), followed by 6 training days where the location of the platform was reversed (reversal trials). Mean±SEM (repeated-measures ANOVA: $f(8.761,714)=1.064$, p=0.388).

Q) The mean number of ultrasonic vocalisations per minute on indicated postnatal days. Mean±SEM (repeated-measures ANOVA: $f(1,72)=0.76$, p=0.39).

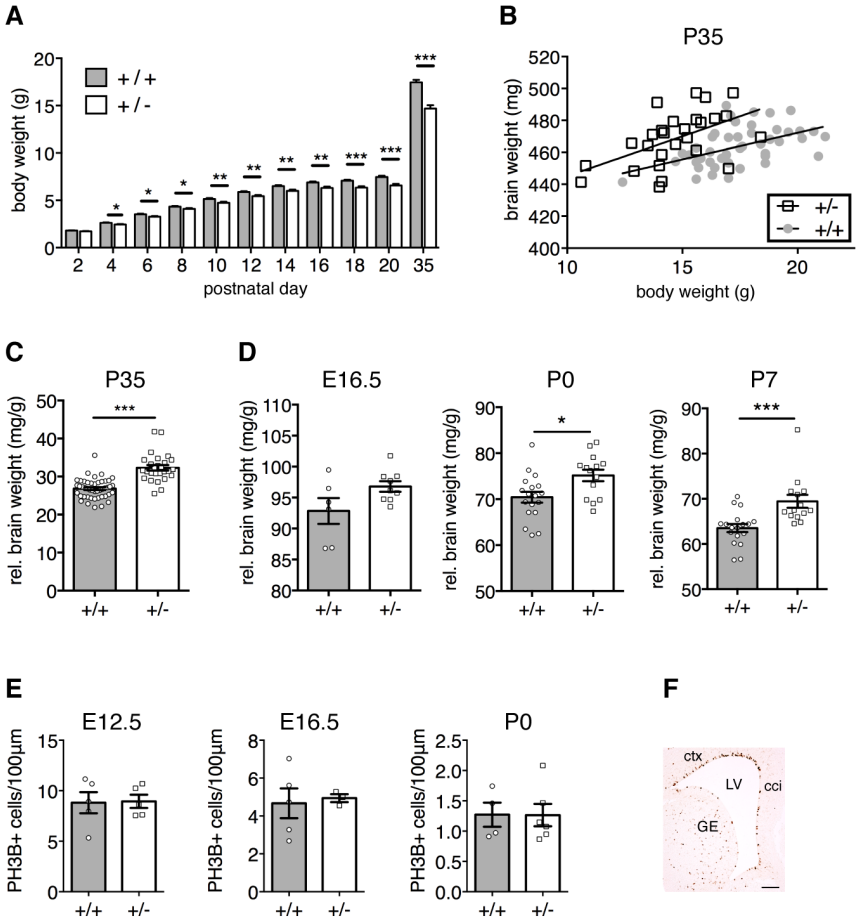


Figure 4. *Chd8* mutants display postnatal brain hypertrophy.

A) Body weights of mice between P2 and P35. Repeated-measures ANOVA with student's t-test as post hoc analysis. ANOVA: f(2.303, 158.940)=12.313, p=0.000003; student's t-tests: P4: t=2.498, *p=0.0148; P6: t=2.385, *p=0.0197; P8: t=1.916, *p=0.0593; P10: t= 2.808, **p=0.0064; P12: t= 2.803, **p=0.0065; P14: t=2.018, **p=0.0035; P16: t=3.353, **p=0.0013; P18: t=4.082, ***p=0.0001; P20: t=4.269, ***p<0.0001; P35: t=6.334, ***p<0.0001; df=71, +/+, n=46; +/-, n=27.

B) Individual wet brain weights plotted against individual body weights for *Chd8*$^{+/-}$ (+/-), mice and their littermate controls (+/+) at postnatal day (P)35. Note that *Chd8*$^{+/-}$ mice have larger brain weights than littermate controls of equivalent body weight. Linear regression analysis: r^2=0.25, p=0.0004 (+/+), r^2=0.28, p=0.005 (+/-).

C) Wet brain weights normalised to body weight P35. *Chd8*$^{+/-}$ (+/-) show significantly increased normalised brain weights compared to their littermate controls (+/+). Mean ± SEM; ***p<0.0001, t=7.455 (student's t-test). +/+, n=46; +/-, n=27.

D) Wet brain weights of pups at P7, P0 and embryonic day (E)16.5 normalised to body weight. *Chd8*$^{+/-}$ (+/-) pups show significantly larger normalised brain weights than their littermate controls (+/+) at P7 and P0. Mean ± SEM; *p=0.01, t-2.746; ***p=0.0009, t=3.681 (student's t-test). P7: +/+, n=18; +/-, n=14; +/+; P0: +/+, n=18; +/-, n=14; +/+; E16.5: +/+, n=6; +/-, n=9.

E) Quantification of phospho-histone H3 (PH3B) positive cells in the ventricular zone at E12.5, E16.5 and P0. Cell counts were normalized to ventricular surface length. Mean ±

SEM; student's t-test. E12.5: +/+, n=5; +/-, n=5. E16.5: +/+, n=5; +/-, n=3. P0: +/+, n=4; +/-, n=6.

F) Example of PH3B immunostaining in an E16.5 coronal brain section. Scale bar=100µm; LV: lateral ventricle, ctx: cortex, cci: cingulate cortex, GE: ganglionic eminence.

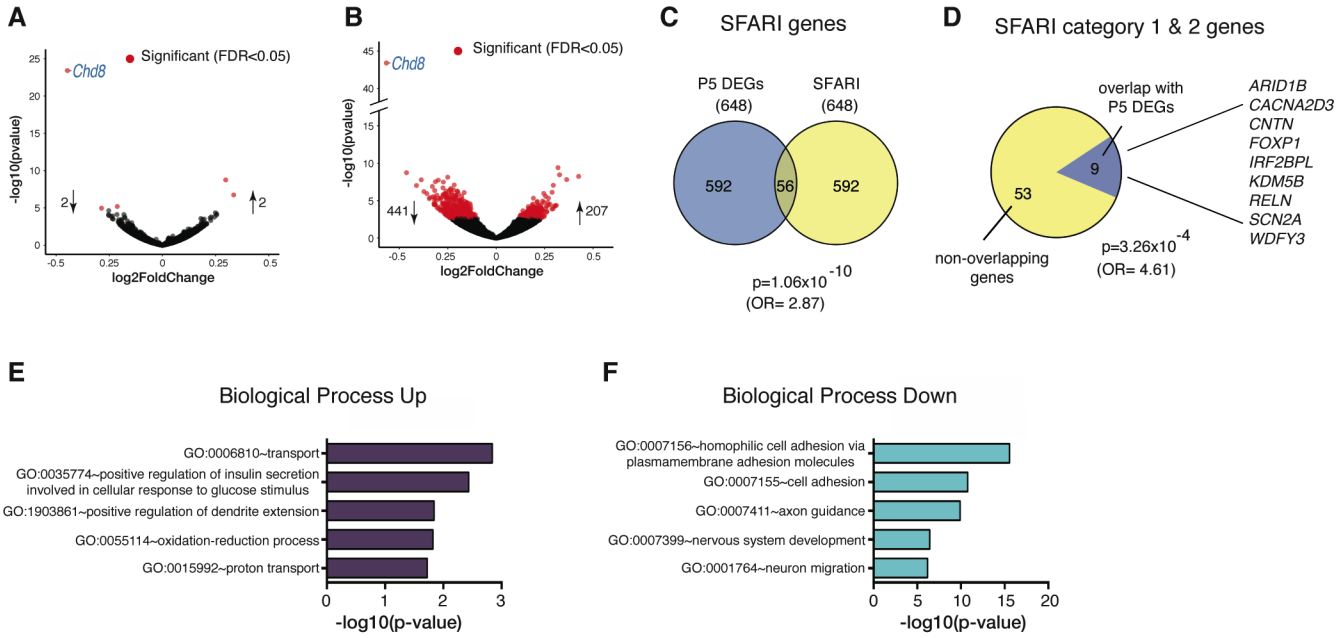


Figure 5. Gene expression changes in *Chd8*-deficient neocortices.
A) Volcano plot of RNA-seq data from embryonic (E12.5) *Chd8*$^{+/-}$ neocortex. Each point represents an individual gene and all genes differentially expressed in *Chd8*$^{+/-}$ samples with an FDR of 0.05 are highlighted in red.
B) Volcano plot indicating differentially expressed genes (DEGs) detected by RNA-seq in P5 *Chd8*$^{+/-}$ neocortex.
C) Venn diagram showing extent of overlap between P5 DEGs and ASD associated genes (categories 1-5 & S) in the SFARI gene database. Enrichment was calculated using Fisher's exact test for count data.
D) Pie chart showing the proportion of high confidence ASD candidate genes (categories 1-2) that are found in the P5 DEG set. Enrichment was calculated using Fisher's exact test for count data.
E) Results of gene set enrichment analysis using the DAVID knowledgebase on the P5 DEG set (FDR<0.05). The five most significant Gene Ontology terms in the Biological Processes category are shown for up-regulated DEGs (left panel) and downregulated DEGs (right panel), respectively.

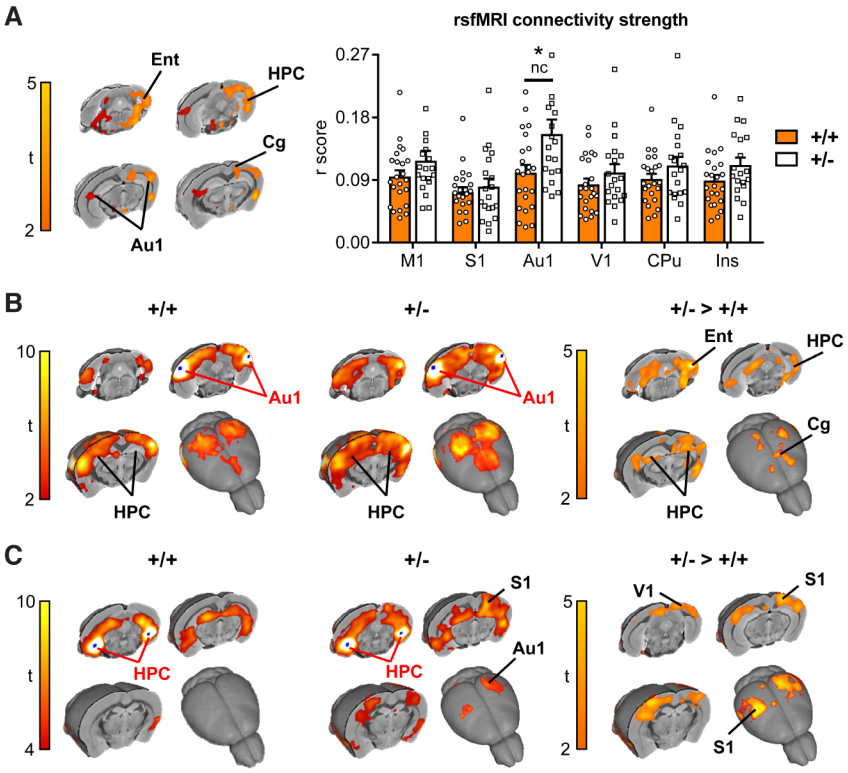
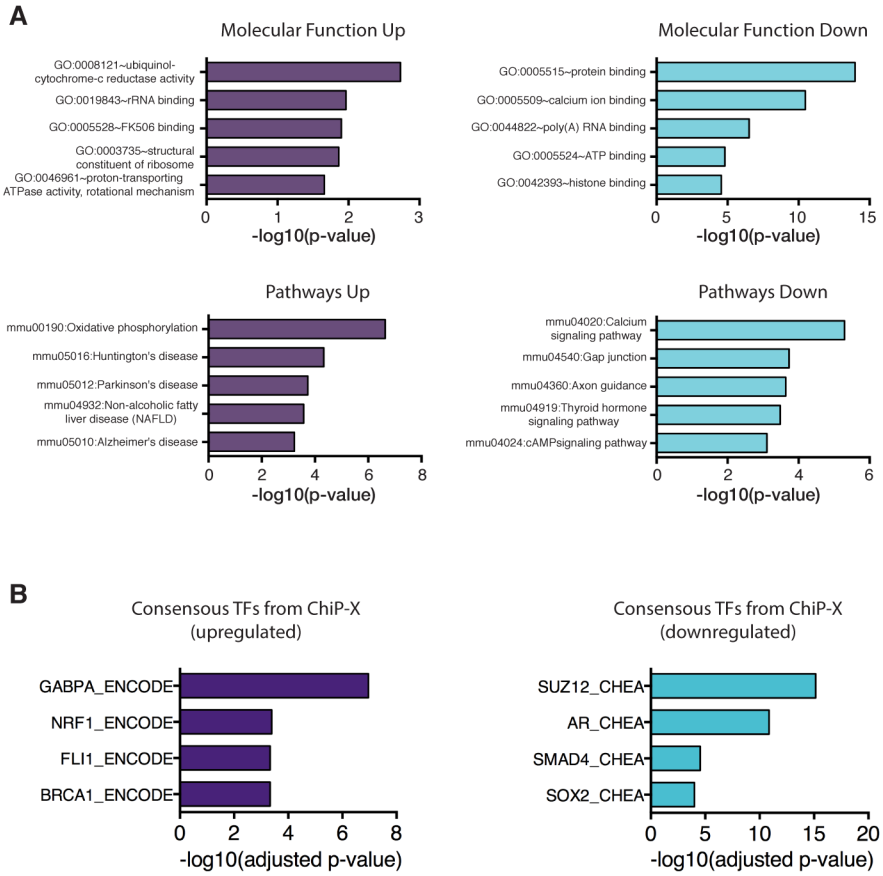


Figure 6: Resting-state functional MRI reveals increased parieto-hippocampal connectivity in *Chd8*$^{+/-}$ mice

A) Global (long range) connectivity mapping revealed bilateral foci of increased connectivity in posterior cortical and hippocampal regions in *Chd8*$^{+/-}$ mice with respect to control littermates (*Chd8*$^{+/-}$: n=19, *Chd8*$^{+/+}$: n=23). The effect is shown at a threshold t_{24}>2.06, P<0.05 (Fig 5A). Orange areas indicate regional rsfMRI differences surviving cluster correction (p=0.05). The bar plot on the right displays global connectivity strength quantification (*r score*) in bilateral areas of interest. *<p0.05, nc=not corrected for multiple comparisons. Ent - entorhinal cortex, HPC - hippocampus, Au1 - auditory cortex, Rs - retrosplenial cortex, M1 - motor cortex, S1 - somatosensory cortex, V1 - visual cortex, CPu - caudate putamen. Ins - insular cortex.

B) Seed based connectivity mapping obtained using a bilateral auditory seed (Au1) covering foci of increased global rsfMRI connectivity depicted in A. A robust increase in rsfMRI connectivity was observed in hippocampal (HPC), entorhinal (Ent) and cingulate (Cg) regions of *Chd8*$^{+/-}$ mice (t<2, pc=0.01).

C) Seed based connectivity mapping obtained using a bilateral ventro-hippocampal seed (HPC) covering bilateral foci of increased global rsfMRI connectivity in A. A significant increase in rsfMRI connectivity was observed in peri-hippocampal and auditory/parietal (S1 and V1) regions in *Chd8*$^{+/-}$ mice (t<2, pc=0.01).



Extended Data Figure 5-1: Functional Enrichment Analysis of differentially expressed genes (DEGs) in *Chd8*$^{+/-}$ P5 neocortices.

A) Results of gene set enrichment analysis using the DAVID knowledgebase on the P5 DEG set (FDR < 0.05). The five most significant Gene Ontology terms in the Molecular Function category are shown for up-regulated DEGs (top left panel) and downregulated DEGs (top right panel), respectively. The five most significantly enriched KEGG pathways are shown for up-regulated DEGs (bottom left panel) and downregulated DEGs (bottom right panel).

B) DEGs (FDR < 0.05) were subjected to enrichment analysis using the "ENCODE and ChEA Consensus TFs from ChIP-X" database on Enrichr, to identify putative upstream regulatory transcription factors. The four most overrepresented transcription factors are shown for up-regulated (left) and down-regulated (right) DEGs, respectively.

Extended Data Tables:

Extended Data Table 2-1: Absolute volumetric differences in specific brain regions between *Chd8*$^{+/-}$ and *Chd8*$^{+/+}$ mice as determined by MRI.

Extended Data Table 5-1: Differentially expressed genes in E12.5 *Chd8*$^{+/-}$ neocortices compared to wildtype controls.

Extended Data Table 5-2: Differentially expressed genes in P5 *Chd8*$^{+/-}$neocortices compared to wildtype controls.

Extended Data Table 5-3: SFARI ASD genes overlapping with P5 differentially expressed genes.

Extended Data Table 5-4: Up-regulated Gene Ontology: Biological Processes

Extended Data Table 5-5: Down-regulated Gene Ontology: Biological Processes

Extended Data Table 5-6: Up-regulated Gene Ontology: Molecular Function

Extended Data Table 5-7: Down-regulated Gene Ontology: Molecular Function

Extended Data Table 5-8: Up-regulated Gene Ontology: Pathways

Extended Data Table 5-9: Down-regulated Gene Ontology: Pathways