CellPress

De novo mutation in genes regulating neural stem cell fate in human congenital hydrocephalus

Charuta G. Furey[1,2*], Jungmin Choi[1*], Sheng Chih Jin[1^], Xue Zeng[1^], Andrew T. Timberlake[1^], Carol Nelson-Williams[1], Mohammad Shahid Mansuri[2], Qiongshi Lu[3], Daniel Duran[1,2], Shreyas Panchagnula[2], August Alloco[2], Jason K. Karimy[2], Jonathan Gaillard[2], Prince Antwi[2], Arjun Khanna[4], Erin Loring[1], William E. Butler[4], Edward R. Smith[5], Benjamin C. Warf[5], David D. Limbrick[6], Phillip B. Storm[7,8], Greg Heuer[7,8], Bermans J. Iskandar[9], James M. Johnston[10], Kaya Bilguvar[1], Shrikant Mane[1], Irina Tikhonova[1], Christopher Castaldi[1], Francesc López-Giráldez[1], James Knight[1], Seth L. Alper[11], Shozeb Haider[12], Bulent Guclu[13], Yasar Bayri[14], Yener Sahin[14], Charles C. Duncan[2], Michael L. DiLuna[2], Murat Günel[1,2], Richard P. Lifton[1,15#], and Kristopher T. Kahle[1,2,16#]

[1]Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA.
[2]Department of Neurosurgery, Yale University School of Medicine, New Haven, Connecticut, USA.
[3]Department of Biostatistics & Medical Informatics, University of Wisconsin, Madison, Wisconsin, USA
[4]Harvard Medical School and Massachusetts General Hospital, Neurosurgical Service, Boston, Massachusetts, USA
[5]Harvard Medical School and Boston Children's Hospital, Boston, Massachusetts, USA
[6]Department of Neurological Surgery and Pediatrics, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA.
[7]Department of Neurosurgery, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[8]Division of Neurosurgery, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.
[9]Department of Neurological Surgery, University of Wisconsin Medical School, Madison, Wisconsin, USA.
[10]Department of Neurosurgery, University of Alabama School of Medicine, Birmingham, Alabama, USA.
[11]Division of Nephrology and Center for Vascular Biology Research, Beth Israel Deaconess Medical Center, Department of Medicine, Harvard Medical School, Boston, MA USA.
[12]Department of Computational Chemistry, University College London School of Pharmacy, London, UK.
[13]Kartal Dr. Lutfi Kirdar Research and Training Hospital, Istanbul, Turkey.
[14]Acibadem Mehmet Ali Aydinlar University, School of Medicine, Department of Neurosurgery, Division of Pediatric Neurosurgery, Istanbul, Turkey.
[15]Laboratory of Human Genetics and Genomics, The Rockefeller University, New York, New York, USA.
[16]Department of Cellular & Molecular Physiology, Yale University School of Medicine, New Haven, Connecticut, USA

* Co-first authors
^ Equal contributors
# Co-corresponding authors

Lead Contact: Kristopher T. Kahle, kristopher.kahle@yale.edu

Abstract

Congenital hydrocephalus (CH), featuring markedly enlarged brain ventricles, is thought to arise from failed cerebrospinal fluid (CSF) homeostasis and is treated with lifelong surgical CSF shunting with substantial morbidity. CH pathogenesis is poorly understood. Exome sequencing of 125 CH trios and 52 additional probands identified three genes with significant burden of rare damaging *de novo* or transmitted mutations: *TRIM71* (p $= 2.15 \times 10^{-7}$), *SMARCC1* (p $= 8.15 \times 10^{-10}$), and *PTCH1* (p $= 1.06 \times 10^{-6}$). Additionally, two *de novo* duplications were identified at the *SHH* locus, encoding the PTCH1 ligand (p $= 1.2 \times 10^{-4}$). Together, these probands account for ~10% of studied cases. Strikingly, all four genes are required for neural tube development and regulate ventricular zone neural stem cell fate. These results implicate impaired neurogenesis and not active CSF accumulation in the pathogenesis of a subset of CH patients, with potential diagnostic, prognostic, and therapeutic ramifications.

Introduction

Hydrocephalus has been defined as the active and progressive distension of the cerebral ventricular system that results from inadequate passage of cerebrospinal fluid (CSF) from its point of production within the cerebral ventricles to its point of absorption into the systemic circulation (Rekate, 2008). This mechanism is most clearly seen in secondary causes of hydrocephalus such as brain tumors, infection, or hemorrhage, in which intracranial pressure is often markedly elevated (Kahle et al., 2016). However, many cases of neonatal or infantile hydrocephalus occur without a known antecedent, and are therefore classified as primary (idiopathic) or congenital hydrocephalus (CH) (Tully and Dobyns, 2014). These congenital forms can occur in the absence of obstruction to CSF flow (communicating hydrocephalus), or with complete/partial intraventricular obstruction (non-communicating hydrocephalus), most often due to aqueductal stenosis. Occasional cases of communicating CH with normal or low intracranial pressure (Bret and Chazal, 1995; Tully and Dobyns, 2014) raise the question of whether the development of increased ventricular volume (i.e., ventriculomegaly) in CH is a primary active process or a secondary passive process.

CH is a major cause of childhood morbidity and mortality, affecting 1 in 1,000 live births (Munch et al., 2012) and representing up to 3% of all pediatric hospital charges in the U.S. (Simon et al., 2008). Accordingly, CH is a major financial burden on health care systems worldwide (Boivin et al., 2015), and costs the U.S health care system alone greater than $2 billion annually (Shannon et al., 2011; Simon et al., 2008). Over the last few decades, there has been little progress in the prevention or treatment of hydrocephalus. Current therapy consists of life-long, catheter-based CSF shunting and endoscopic third ventriculostomy with or without choroid plexus cauterization, invasive surgeries with high rates of failure and morbidity (Kahle et al., 2016). Our lack of understanding of the molecular pathogenesis of CH constitutes a fundamental barrier to developing improved approaches to diagnosis and treatment (McAllister et al., 2015).

Despite several reports of multiplex pedigrees that feature hydrocephalus as the predominant phenotypic feature, few *bona fide* human CH-causing genes have been identified (Kousi and Katsanis, 2016). These include *L1CAM* mutation in X-linked hydrocephalus and aqueductal stenosis (OMIM#307000; (Rosenthal et al., 1992)), which constitutes up to 3% of CH, and rare recessive mutations in *MPDZ* (OMIM#603785; (Al-Dosari et al., 2012)), *CCDC88C* (OMIM#236600; (Ekici et al., 2010)), *EML1* and *WDR81* (Shaheen et al., 2017). These genes appear to affect multiple cellular processes, including neural cell adhesion, planar cell polarity, and cellular vesicle transport, confounding efforts to formulate a uniform paradigm of CH pathophysiology (Kousi and Katsanis, 2016). Although more than 40% of all CH cases are predicted to have a genetic etiology (Zhang et al., 2006), mutations in currently identified genes account for <5% of primary CH cases (Adle-Biassette et al., 2013; Haverkamp et al., 1999).

The sporadic nature of most CH cases, which present as a single affected subject in an otherwise unaffected family, has limited the utility of traditional genetic approaches to their investigation. This limitation has motivated whole-exome sequencing (WES) in large numbers of such subjects, searching for genes mutated in affected subjects more often than expected by chance. This has proven a powerful approach in the study of genetically heterogeneous

neurodevelopmental disorders (Deciphering Developmental Disorders Study, 2015, 2017), including brain malformations (Barak et al., 2011; Bilguvar et al., 2010; Mishra-Gorur et al., 2014), epilepsy (Allen et al., 2013), intellectual disability (Gilissen et al., 2014), craniosynostosis (Timberlake et al., 2016; Timberlake et al., 2017), schizophrenia (Xu et al., 2011), autism (Awadalla et al., 2010; Iossifov et al., 2012; Krumm et al., 2015; Neale et al., 2012; O'Roak et al., 2011; O'Roak et al., 2012; Sanders et al., 2012), obsessive compulsive disorder (Cappi et al., 2016; Cappi et al., 2017), attention-deficit/hyperactivity disorder (Satterstrom et al., 2018), and Tourette's syndrome (Willsey et al., 2017). We hypothesized that the apparent sporadic occurrence of CH might in some cases reflect damaging *de novo* mutations and/or transmitted mutations with incomplete penetrance.

Results

Cohort characteristics and whole exome sequencing

We recruited 177 probands with non-*L1CAM*-mutated primary CH in which hydrocephalus was the predominant phenotypic feature (see **Methods**). All probands had undergone surgery (shunt placement and/or endoscopic third ventriculostomy) for CSF diversion. The cohort included 125 parent-offspring trios with a single affected offspring, 47 singleton cases, and 5 multiplex kindreds (see **Methods**; **Supplementary Table 1**). 88 probands had communicating hydrocephalus and 89 had aqueductal stenosis.

DNA was isolated and WES was performed as described (Timberlake et al., 2016). 95.9% of targeted bases had 8 or more independent reads, and 92.7% had 15 or more (**Supplementary Table 3**). 1,789 control trios that comprised unaffected siblings from the Simons simplex autism cohort were sequenced on the same platform to a similar depth of coverage and were analyzed in parallel (Krumm et al., 2015). In both cohorts, variants were called using the Genome Analysis Toolkit (GATK) Haplotype Caller (McKenna et al., 2010; Van der Auwera et al., 2013) and allele frequencies were annotated in the Exome Aggregation Consortium (ExAC) and gnomAD databases ((Lek et al., 2016); see **Methods**). TrioDeNovo was used to identify *de novo* mutations (Dong et al., 2015) and MetaSVM was used to infer the impact of missense mutations (Dong et al., 2015). Direct Sanger sequencing of PCR amplicons containing the mutation verified mutations in genes of interest.

We identified in 4 probands mutations in genes previously associated with isolated or syndromic inherited human CH, including one proband who was compound heterozygous for loss-of-function mutations in *MPDZ* (Hydro182; (Al-Dosari et al., 2012)); two male probands with transmitted damaging missense mutations in the X-linked gene *FLNA*, known to cause the neuronal migration disorder, periventricular nodular heterotopia (Hydro131 and Hydro169; OMIM#300049; (Kamuro and Tenokuchi, 1993; Sheen et al., 2004)); and a single proband with a compound heterozygous mutation in the crumbs cell polarity complex *CRB2* (Hydro122), known to cause a cilia-associated syndrome of cystic kidney disease and ventriculomegaly (OMIM# 219730; (Slavotinek et al., 2015)).

Global analysis of burden of *de novo* mutation

We determined a *de novo* mutation rate of 1.4×10^{-8} per base pair, with 1.09 *de novo* coding region mutations per proband (**Table 1**). These results are consistent with both expectation and prior experimental results (Homsy et al., 2015; Timberlake et al., 2017; Ware et al., 2015). The burden of *de novo* mutations in the control cohort was similar (**Supplementary Table 4**).

We compared the observed and expected number of *de novo* mutations in all genes and among genes in the top quartile of brain expression at embryonic (E) day 9.5 (see **Methods**). Protein-altering *de novo* mutations were significantly enriched over expectation in the latter group, contributing to an estimated 8% of cases (**Table 1**). In contrast, controls showed no significant enrichment of *de novo* mutations in any gene class (**Supplementary Table 4**).

We identified five genes with two or more protein-altering *de novo* mutations: *TRIM71*, *SMARCC1*, *PTCH1*, *PLOD2*, and *SGSM3*. The first three of these genes have a probability of loss-of-function intolerance (pLI) score ≥ 0.99, while the latter two have pLIs of zero (Lek et al., 2016). The probability of finding five genes with two or more protein-altering *de novo* mutations in a cohort of this size by chance alone is very low ($p = 1.1 \times 10^{-4}$). Four of these genes were in the upper quartile of expression in developing brain. The probability of 2 or more protein-altering *de novo* mutations in 4 such genes in our cohort is very low ($p = 1.1 \times 10^{-4}$; **Supplementary Table 6**).

Recurrent *de novo* and transmitted mutations in *TRIM71*

LIN41/*TRIM71*, encoding lineage variant 41 (LIN41)/tripartite motif 71 (herein termed "TRIM71") (Reinhart, 2000; Slack, 2000), harbored three novel *de novo* mutations, including the identical heterozygous *de novo* p.Arg608His mutation in two unrelated probands with severe, prenatally diagnosed, non-obstructive (i.e., communicating) hydrocephalus (Hydro101-1 and Hydro102-1) (**Figure 1**; **Supplementary Table 7**). During the course of this study, Hydro102-1 became pregnant and transmitted the p.Arg608His mutation to her son, who was diagnosed with severe communicating hydrocephalus on prenatal ultrasound (**Figure 1**). A third *de novo* *TRIM71* mutation (p.Arg796His) was identified in Hydro100-1, who had prenatally diagnosed communicating CH (Hydro100-1). All CH patients harboring *TRIM71* mutations underwent surgical CSF shunting at birth. The probability of finding three or more protein-altering *de novo* mutations in *TRIM71* by chance in a cohort of this size is 2.15×10^{-7}, surpassing genome-wide significance (**Table 2**; (Ware et al., 2015)). *TRIM71* not only has a pLI of 0.99, but is also exceptionally intolerant to missense variation (z score = 5.69) (Lek et al., 2016). Moreover, with 105 protein-altering *de novo* missense mutations in our cohort, the probability of seeing any instances of the identical mutation at any position in the coding region by chance is very remote ($p = 5.24 \times 10^{-4}$).

TRIM71 mRNA is the closest human homolog of *C. elegans Lin-41*, the evolutionarily conserved direct target of let-7 (lethal 7) microRNA (miRNA) in the heterochronic pathway regulating developmental transitions between stem cell proliferation and differentiation (Ecsedi and Grosshans, 2013). TRIM71, like Lin-41, mediates post-transcriptional silencing of mRNAs via direct interactions of its NHL domain with specific sequences in the 5' or 3' UTRs of target genes (Aeschimann et al.; Ecsedi and Grosshans, 2013; Slack and Ruvkun, 1998; Vella et al., 2004), and also contains ubiquitin ligase activity conferred by its RING domain (Nguyen et al.,

2017). The NHL domain in TRIM71 consists of 6 repeats, each 40-50 residues long, that jointly comprise a barrel-like six-bladed β-propeller (Loedige et al., 2015); the *de novo* p.Arg608His and p.Arg796His mutations are at homologous positions in different blades of the NHL domain. Arginines at these positions are completely conserved among orthologs from *H. sapiens* to *C. elegans*, and occur in a highly conserved RPQGV motif (**Figure 1**). The NHL domain and RNA binding motif of TRIM71 is the closest homolog of the NHL domain of *D. melanogaster* Brat (Loedige et al., 2013; Loedige et al., 2014), which has been co-crystalized bound to a target RNA (Loedige et al., 2015). In this structure, the amino groups of side chains of the mutated arginines form hydrogen bonds with either the phosphate backbone (p.Arg796) or a uracil base (p.Arg608) of target RNA (**Figure 1**; **Supplementary Figures 4-7**); these interactions are predicted to be altered by histidine substitution. While these mutations were predicted as tolerated by MetaSVM, they are predicted as damaging by 43 of 45 other predicting algorithms, including the very conservative MPC-D algorithm (see **Methods;** (Samocha et al., 2017)). The recurrence of the identical *de novo* mutation and the fact that the third *de novo* mutation is at a homologous position in a different blade of the NHL propeller domain strongly suggests that these mutations in *TRIM71* are not simple loss-of-function (LOF) mutations.

In situ hybridization in embryonic day 12.5 (E12.5) mouse brain (**Figure 1; Supplementary Figure 8**) revealed abundant *Trim71* expression in the ciliated neuroepithelium and ventricular zone (Cuevas et al., 2015; Maller Schulman et al., 2008). Analogous to its heterochronic expression in *C. elegans* (Kanamoto, 2006; Slack, 2000), *Trim71* expression significantly decreases during development (**Figure 1**) (Schulman et al., 2005; Yu et al., 2010), with gradual restriction of transcription to neural tissues and limb buds (Chen et al., 2012; Maller Schulman et al., 2008). Trim71 is highly expressed in neural progenitor cells of early mouse embryos (until E11.5) but declines secondary to increased expression of let-7 and mir-125 miRNAs as neural differentiation proceeds (Chen et al., 2012; Schulman et al., 2005). Global Trim71 deletion in mice results in early embryonic lethality, with severe exencephaly (failure of closure of the cephalic end of the neural tube) by E10 (Maller Schulman et al., 2008). Trim71 maintains the pluripotency of neural progenitor cells by regulating the balance between self-renewal and differentiation via the post-transcriptional silencing of its target mRNAs (Chang et al., 2012; Ecsedi and Grosshans, 2013; Mitschka et al., 2015; Worringer et al., 2014). The neural tube closure defect in *Trim71* knockout mice results from decreased proliferation and precocious differentiation of neural progenitor cells (Chen et al., 2012).

Multiple *de novo* and transmitted mutations in *SMARCC1*

Two novel, damaging *de novo* mutations were identified in *SMARCC1*, encoding BAF155, a 155-kD core subunit of the mammalian B̲RG1/BRM a̲ssociated f̲actor (BAF; *S. cerevisiae* SWI/SNF) chromatin remodeling complex (Wang et al., 1996). These mutations included a p.Lys891fs*4 mutation (in Hydro106-3) and a predicted damaging missense p.His526Pro mutation (in Hydro105-1) located in an entirely conserved position in the SWIRM domain, which mediates protein-protein interactions with other BAF complex subunits (Da et al., 2006) (**Figure 2**). p.His526Pro occurred in a proband with severe obstructive CH with aqueductal stenosis and no other affected family members. p.Lys891fs*4 was *de novo* in an unaffected father (Hydro106-3) of three children with severe, prenatally-diagnosed CH with aqueductal stenosis. Two of these offspring died *in utero* from severe obstructive hydrocephalus, while the

surviving third child inherited p.Lys891fs*4. In this cohort, the probability of seeing two damaging *de novo* mutations in *SMARCC1* was 2.69 x 10⁻⁶ (**Table 2**).

Three additional, previously unidentified, transmitted LOF mutations in *SMARCC1* were found in three other CH probands presenting with severe obstructive CH with aqueductal stenosis. In each of these kindreds, the proband was the sole affected member with transmission of the mutation from an unaffected parent (**Figure 1**; **Supplementary Table 8**). The probability of two *de novo* damaging mutations and three rare transmitted LOFs occurring in *SMARCC1* in this cohort was 8.15 x 10⁻¹⁰; this is a conservative estimate since *SMARCC1* is highly intolerant to mutation (pLI of 1), with far fewer LOF mutations than expected (3 LOF *SMARCC1* mutations among 60,706 unrelated individuals in the ExAC database, and 2 among 3,578 autism parental controls).

In situ hybridization demonstrated that *Smarcc1*, like *Trim71*, is highly expressed in the ciliated neuroepithelium and ventricular zone, but significantly reduced in later development and in adult brain (Ho et al., 2009; Tuoc et al., 2013a; Yan et al., 2008) (**Figure 2; Supplementary Figure 8**). Homozygous null mutation of *Smarcc1* results in early embryonic lethality soon after decidualization due to defects in the inner cell mass (Kim et al., 2001). ~20% of heterozygous null mutants (Kim et al., 2001) and ~80% of mice homozygous for a missense allele (*Baf155^{msp/msp}*) (Harmacek et al., 2014) develop exencephaly similar to that of *Trim71* knockout mice (Harmacek et al., 2014; Narayanan et al., 2015; Nguyen et al., 2016). The exencephaly phenotype has been associated with inappropriate proliferation and increased apoptosis of neuroprogenitor cells in the neural tube (Harmacek et al., 2014; Kim et al., 2001). A conditional neuronal knockout of mouse *Smarca4*, encoding a binding partner of Smarcc1 in the BAF complex, causes severe hydrocephalus, aqueductal stenosis, and thinning of the cerebral cortex (Cao and Wu, 2015). Neural progenitor cell-specific BAF complexes that include SMARCC1 regulate proliferation, differentiation, and survival of mouse neural progenitor cells via ATP-dependent remodeling of chromatin and associated transcriptional regulation of genes critical for neurogenesis during telencephalon development (Narayanan et al., 2015).

De novo and transmitted mutations in *PTCH1*

Two previously unidentified *de novo* LOF mutations in CH probands were identified in *PTCH1*, encoding Patched-1 (PTCH1), the transmembrane receptor for Sonic Hedgehog (SHH) in primary cilia (Eggenschwiler and Anderson, 2007) (**Figure 3**). Heterozygous LOF mutations in PTCH1 have previously been implicated in Gorlin syndrome (OMIM#109400), featuring numerous basal cell carcinomas and frequent tumors of the mandible, along with variable expressivity of other skeletal and non-skeletal features (8681379). *De novo* mutations in CH patients included a start-loss mutation (p.Met152fs*1) and a splice donor site mutation (c.1503+3T>C) in unrelated probands presenting with severe CH and aqueductal stenosis (**Figure 3**; **Supplementary Table 9**). Both mutations were heterozygous and occurred in probands of uniplex CH kindreds. A transmitted frameshift mutation in *PTCH1* (p.Leu664fs*12) was also found in two brothers with severe CH (**Figure 3**; **Supplementary Table 9**). This mutation was transmitted from a mother who did not have hydrocephalus but had been diagnosed with Gorlin syndrome on the basis of numerous basal cell carcinomas. The probability of seeing

at least two damaging *de novo* mutations by chance in a cohort of this size is 1.06×10^{-6}, surpassing genome-wide significance (**Table 2**; (Ware et al., 2015)).

Consistent with previous results (Eggenschwiler et al., 2001; Goodrich et al., 1996; Goodrich et al., 1997; Takahashi and Osumi, 2002), *in situ* hybridization showed high and specific expression of *Ptch1* in hindbrain neuroepithelium (**Figure 3; Supplementary Figure 8**). Global knockout of *Ptch1* in mouse results in lethality between E9.0 and E10.5, with exencephaly and neural tube overgrowth (Ellis et al., 2003; Goodrich et al., 1997; Milenkovic et al., 1999). A significant fraction of *Ptch1*$^{+/-}$ mice develop severe hydrocephalus in two different genetic backgrounds (Svärd et al., 2009; Wetmore et al., 2000), similar to the variable expressivity seen in one of the multiplex kindreds. Penetrance of hydrocephalus in these mice increases to 100% in conjunction with the homozygous quaking viable mutation (*Ptch1*$^{+/-}$; $qk^{v/v}$) (Gavino and Richard, 2011). Primary cilia in neuroepithelial cells sense gradients of SHH via PTCH1 and transduce these signals to regulate neural progenitor cell proliferation, differentiation, and fate specification in the developing and adult mouse CNS (Palma et al., 2005; Palma and Ruiz i Altaba, 2004).

De novo mutations in *PLOD2*

Two heterozygous *de novo* missense mutations were identified in the highly brain expressed (HBE) gene *PLOD2*, which has never before been implicated in human hydrocephalus (**Supplementary Tables 10 and 11**). *PLOD2* did not surpass thresholds for genome-wide significance, and mutations in *PLOD2* were predicted to be damaging by MetaSVM. *PLOD2* encodes a lysyl hydroxylase previously implicated in autosomal recessive Bruck syndrome type 2 (OMIM#609220), a variant of osteogenesis imperfecta (McPherson and Clemens, 1997). Further work will be required to assess the significance of these mutations in CH.

De novo mutations in neural tube closure and formation

GO pathway analysis of our significant gene set identified them to be highly enriched for "neural tube closure and formation" (**Supplementary Table 12**). We consequently searched for additional *de novo* mutations in CH genes in this pathway (**Supplementary Table 13**). We found a single *de novo* missense mutation in *CELSR2* (p.R2812W), encoding Cadherin EGF LAG Seven-Pass G-Type Receptor 2, and a single *de novo* missense mutation in *LRP6* (p.V1415F), encoding a co-receptor with frizzled proteins in the WNT signaling pathway. Knockout of each of these genes in mice results in progressive and lethal hydrocephalus (Allache et al., 2014; Cuevas et al., 2015; Gavino and Richard, 2011; Harmacek et al., 2014; Maller Schulman et al., 2008; Tissir et al., 2010). Both *CELSR2* and *LRP6* are highly expressed in the neuroepithelium and ependyma, and regulate neurogenesis. These observations suggest that the identified mutations in *CELSR2* and *LRP6* contribute to human CH pathogenesis.

De novo duplications at the *SHH* locus

Using the XHMM algorithm (Fromer and Purcell, 2001), we identified seven putative *de novo* copy number variants (CNVs) see **Methods**). Six of these, including five duplications and one deletion, were validated by independent tests (distortion of allelic ratios of heterozygous SNPs in

implicated intervals or qPCR; (**Supplementary Table 14**). Two of these were duplications at the *SHH* locus, encoding Sonic Hedgehog (SHH), the canonical ligand for PTCH1 that regulates neurogenesis by conferring positional information to ventral neural progenitors in the developing neural tube (see Jessell, 2000; Lupo et al., 2006). The probability of finding two duplications at the same locus among 5 duplication events with these gene compositions is very low (p = 1.22 x 10^{-4}) (see **Methods**).

Discussion

Using exome sequencing, we implicate four new genes in the molecular pathogenesis of human CH: *TRIM71*, *SMARCC1*, *PTCH1*, and *SHH*. High enrichment of mutations in these genes among our CH cohort suggests their large effects on phenotypic risk. The genes showing heterozygous protein-altering *de novo* mutations are all highly intolerant to mutation (pLI > .99), and murine knockouts of each gene produce severe defects in neural tube development, including exencephaly (Feng et al., 2013; Gavino and Richard, 2011; Goodrich et al., 1997; Hahn et al., 1996; Harmacek et al., 2014; Maller Schulman et al., 2008). Two of these genes encode proteins that are binding partners: SHH is the canonical ligand for PTCH1. A subset of mice deficient in *ptch1* develops fatal hydrocephalus (Celen et al., 2017; Gavino and Richard, 2011). All four genes are highly expressed in the developing ciliated neuroepithelium, and regulate neural stem cell fate. Together, mutations in these four new genes collectively account for 8.5% of cases studied, with another 2.3% of cases explained by previously identified genes.

From the inferred contribution of *de novo* point mutations to 8% of CH probands in our cohort, and the number of genes with more than one *de novo* mutation, we estimate that a small number of additional genes contribute to CH by *de novo* point mutation (maximum likelihood estimate ~8 genes, with wide confidence interval) (**Supplementary Figures 13 and 14).** This estimate is substantially smaller than found for other diseases such as autism and congenital heart disease (Homsy et al., 2015; Iossifov et al., 2014; Sanders et al., 2012; Zaidi et al., 2013; De Rubeis et al., 2014), and is consistent with the identification of multiple disease-causing genes from analysis of this relatively small cohort. Variants in non-coding elements of these genes might add to their contribution to CH. Sequencing of additional trios and isolated probands has high potential to detect additional rare mutations with large effect on disease risk.

Interestingly, all four cases of CH with *TRIM71* mutation had communicating hydrocephalus, whereas aqueductal stenosis characterized all five cases with *SMARCC1* mutation and all three cases with *PTCH1* mutation. From the prevalence of these two sub-phenotypes in our cohort (50.2% communicating vs. 49.8% aqueductal, respectively) we calculate that this correlation was highly unlikely to occur by chance (P = 0.002), supporting the pathogenicity of the observed mutations. These observations provide evidence that phenotypic subsets of CH may be strongly influenced by genetic determinants.

Nonetheless, our results also provide evidence of incomplete penetrance and variable expressivity of mutations in two of these genes. Indeed, there were four first-degree relatives of probands with the same mutation found in the proband in *SMARCC1* and one relative with *PTCH1* mutation who had no history or evidence of hydrocephalus, demonstrating incomplete penetrance for these mutations. Moreover, heterozygous LOF mutations in *PTCH1* are a well-

described cause of Gorlin syndrome (Gorlin, 2004; Gorlin and Goltz 1960); Gorlin syndrome patients typically only develop hydrocephalus with development of obstruction to CSF flow due to medulloblastoma (Amlashi et al., 2003). In contrast, our patients had no obstruction and did not have medulloblastoma; murine hydrocephalus with PTCH1 mutation similarly occurs without medulloblastoma (Wang et al., 2010a; Wetmore et al., 2015). Moreover, a mutation-bearing parent of one of our probands had typical Gorlin syndrome diagnosed owing to numerous basal cell carcinomas, but did not have hydrocephalus. This provides strong support that the LOF mutations in our probands are functionally equivalent to those causing Gorlin syndrome and implies that these probands are at high risk to develop basal cell carcinomas later in life. The determinants of incomplete penetrance and variable expressivity remain obscure, but may be explained by common genetic (Timberlake et al., 2016) or environmental modifiers (Stuart et al., 2015), or by stochastic elements.

Similar to other Mendelian forms of CH (e.g., *L1CAM*) (Rosenthal et al., 1992), several CH patients harboring *TRIM71* and *SMARCC1* mutations exhibited other neurological and brain radiographic features in addition to the predominant phenotype of hydrocephalus. For example, all patients with *TRIM71* mutations, and 3 of 5 patients with SMARCC1 mutations, exhibited mild to severe global neurodevelopmental delay and epilepsy. A mother and son with mutant *TRIM71* and ventriculomegaly had similar open schizencephalic clefts (**Figure 1**). In addition, the mother had polydactyly and her son had bilateral radial aplasia. Two patients with *SMARCC1* mutations also appeared to have other structural brain abnormalities that included open schizencephalic clefts (**Figure 2**). These findings are consistent with known roles of TRIM71 and SMARCC1 in regulating key aspects of early brain development (Cuevas et al., 2015; Harmacek et al., 2014; Kim et al., 2001; Maller Schulman et al., 2008; Narayanan et al., 2015; Nguyen et al., 2016).

While mutations in *SMARCC1* and *PTCH1* in CH are LOF mutations, this is not likely the case for *TRIM71*, which contains three *de novo* mutations, all of which involve either of two homologous arginines in the NHL domain that directly bind to target mRNAs (Chang et al., 2012; Loedige et al., 2013; Loedige et al., 2015; Loedige et al., 2014). This highly specific mutation spectrum strongly suggests that these are not simple LOF mutations. While these mutations likely impair binding to normal target mRNAs, there may be other gene functions that are preserved; for example, the putative ubiquitin ligase activity of TRIM71 might serve an independent function that is preserved. Alternatively, the mutations could have neomorphic effects, for example by binding to mRNAs that are not normally targeted. These issues will be interesting topics for future investigation. Similarly, the observed duplications of *SHH* suggest that these mutations are gain of function. Dosage loss of PTCH1, which inhibits downstream signaling in the pathway via effect on the *SMO* receptor, and dosage gain of *SHH*, are both expected to increase downstream signaling in the hedgehog pathway (Briscoe and Therond, 2013).

Ciliated neuroepithelial stem cells (i.e., neural progenitor cells, or neural stem cells) in the ventricular zone of the neural tube undergo proliferation accompanied by minimal differentiation during early embryogenesis. These cells subsequently become differentiated neurons (Ever and Gaiano, 2005; Merkle et al., 2004) in mice between E8-14 (Chen et al., 2006; Noctor et al., 2004), a transformation driven by significant alterations in the expression of critical

developmental genes (Yao et al., 2016). TRIM71, SMARCC1, and PTCH1 are all highly expressed in the developing neuroepithelium (Cuevas et al., 2015; Gavino and Richard, 2011; Tuoc et al., 2013b). The secreted factor SHH engages its receptor PTCH1 on neuroepithelial cilia. All 5 genes have been shown to regulate the proliferation, fate specification, and differentiation of neural progenitor cells, and accordingly, the development of the neural tube (Kim et al., 2001; Nguyen et al., 2017; Shikata et al., 2011).

Although CH research has historically focused on disordered CSF homeostasis (Oreskovic and Klarica, 2011; Sato et al., 1994), the CH gene mutations reported here suggest the importance of *de novo* or transmitted mutations that alter the balance between neural progenitor cell proliferation and neural differentiation. In support of this human genetic data, *Dusp16* deficiency in mice causes brain overgrowth and associated obstructive hydrocephalus due to hyperproliferation and expansion of ventricular zone neural progenitor cells at the level of the aqueduct (Zega et al., 2017). In contrast, the communicating hydrocephalus phenotype in a mouse model of the human ciliopathy Bardet-Biedl Syndrome results from decreased proliferation and survival of ventricular zone neural progenitor cells (Carter et al., 2012). Interestingly, intraventricular hemorrhage in neonates, the most common cause of secondary hydrocephalus in children (Mazzola et al., 2014; Robinson, 2012), also results from neuronal deficiency due to ependymal denudation, decreased neural progenitor cell viability, and associated cortical thinning (McAllister et al., 2017; Yung et al., 2011). Therefore, a "neural stem cell" paradigm of hydrocephalus may provide a unifying mechanism to explain multiple forms of neonatal hydrocephalus.

Lastly, these findings have implications for our understanding of CH treatment. Multiple non-neural mechanisms have been proposed to account for idiopathic communicating hydrocephalus, including elevated venous pressure, arachnoid granulation immaturity, excessive skull growth, and lymphatic dysplasia (Govaert et al., 1991). However, given the new genetic data presented herein, it seems that a primary driver of some communicating forms of CH, especially in those associated with other structural brain abnormalities, may not be attributable to active CSF accumulation and associated ventricular distension, but rather impaired neurogenesis. In neonatal CH cases with normal or low intracranial pressure (Tully et al., 2016), it is unclear if surgical CSF shunting addresses a pathologic feature of CH or, rather, only exposes CH patients to the morbidity of surgery and life-long shunt complications without significantly improving neurodevelopmental outcomes. The direct and obvious implications for families with CH children and their neurosurgeons will stimulate discussion and catalyze further investigation into CH pathogenesis and treatment.

Acknowledgements

Author Contributions

C.G.F., R.P.L, and K.T.K conceived, designed and directed study; C.G.F, D.D., S.P, A.A., J.K., J.G., P.A., A.K., E.L., W.E.B., E.R.B., B.C.W., D.D.L., P.B.S., G.H., B.J.I., J.M.J., B.G., Y.B., Y.S., C.C.D, M.L.D. performed clinical evaluations and enrolled patients; C.G.F., A.T.T., C.N.W., M.S.W., K.B., S.M., I.T., C.C., F.L.G, J.K., S.L.A., S.H. performed research; C.G.F., J.C., S.C.J., X.Z., A.T.T., Q.L., R.P.L, K.T.K. analyzed data; and C.G.F., R.P.L, and K.T.K. wrote the manuscript.

STAR Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Kristopher T. Kahle (kristopher.kahle@yale.edu).

Experimental Model and Subject Details

Subjects and samples

All study procedures and protocols comply with Yale University's Human Investigation Committee and Human Research Protection Program. Written informed consent for genetic studies was obtained from all participants. Inclusion criteria included patients with primary congenital hydrocephalus without known genetic causes such as *L1CAM* or any other large chromosomal deletions or rearrangements. Hydrocephalus cases with secondarily acquired etiologies such as intraventricular hemorrhage (IVH), meningitis, obstruction due to tumors or cysts, and stroke were excluded. Children with hydranencephaly, large cysts and cephaloceles, posterior fossa crowding, myelomeningocele (Chiari II syndrome), or benign extra-axial CSF accumulation (i.e. benign external hydrocephalus) were also excluded. Sequenced trios were composed of two unaffected parents and one affected child with primary congenital hydrocephalus. All probands had undergone surgery for therapeutic CSF diversion (shunt placement and/or endoscopic third ventriculostomy). Patients and participating family members provided buccal swab samples (Isohelix SK-2S DNA buccal swab kits), medical records, radiological imaging studies, operative reports, and congenital hydrocephalus phenotype data.

The control cohort was composed of 1,789 previously whole-exome sequenced and analyzed families from the Simons Foundation Autism Research Initiative Simplex Collection (Fischbach and Lord, 2010; Iossifov et al., 2014; Krumm et al., 2015; O'Roak et al., 2011; Sanders et al., 2012). Sequenced families were comprised of two unaffected parents, one affected child with autism, and one unaffected sibling. Only the unaffected sibling and parents, as designated by the Simons Simplex Collection, were analyzed and served as controls for this study (Krumm et al., 2015).

Whole-exome sequencing and variant calling

DNA was isolated from buccal swab samples in accordance with manufacturer protocol. Whole-exome sequencing was performed using the IDT xGen capture reagent followed by 99 base paired-end sequencing on the Illumina HiSeq 2000 instrument at the Yale Center for Genome Analysis as previously described (Timberlake et al., 2016). Exome sequencing quality metrics are shown in **Supplementary Table 3**.

Sequence reads were mapped and aligned to the GRCh37/hg19 human reference genome using Burrows-Wheeler Aligner-MEM (Li and Durbin, 2009). In accordance with GATK Best Practices recommendations, the data were further processed using Genome Analysis Toolkit (GATK) base quality score recalibration (McKenna et al., 2010), indel realignment, duplication

marking and removal, and base quality score recalibration (DePristo et al., 2011; Van der Auwera et al., 2013). Single nucleotide variants and small insertions and deletion were called using GATK Haplotype Caller and annotated using ANNOVAR (Wang et al., 2010b), NHLBI exome variant server (Fu et al., 2013), 1000 Genomes (The Genomes Project, 2015), DbSNP (Sherry et al., 2001), and gnomAD and ExAC databases (Lek et al., 2016).

The average depth of coverage of the whole-exome sequencing data was 54.3x, with greater than 8x coverage in 95.9% of the target region for exome capture (**Supplementary Table 3**).

The sporadic or autosomal recessive mode of inheritance exhibited in our cohort pedigrees led us to prioritize *de novo*, compound heterozygous, and homozygous variants. Variants were filtered for predicted deleteriousness and conservation using a series of *in silico* prediction algorithms, including Meta-analytic support vector machine (MetaSVM) and Meta-analytic logistic regression (MetaLR) (Dong et al., 2015), Polymorphism Phenotyping (PolyPhen) (Adzhubei et al., 2010), Combined Annotation-Dependent Depletion (CADD) (Kircher et al., 2014), Sorting Intolerant From Tolerant (SIFT) (Kumar et al., 2009; Ng and Henikoff, 2001, 2002, 2003, 2006), conservation across 46 orthologs (cons46diff) (Cromer et al., 2012; Stuart et al., 2015), Likelihood Ratio Test (LRT) (Chun and Fay, 2009), MutationTaster (Schwarz et al., 2010), Mutation Assessor (Reva et al., 2011), Functional Analysis Through Hidden Markov Models (FATHMM) (Shihab et al., 2013), FATHMM–Multiple Kernel Learning (FATHMM-MKL) (Shihab et al., 2015), FATHMM-Coding (Shihab et al., 2014), Protein Variant Effect Analyzer (PROVEAN) (Choi et al., 2012), Variant Effect Scoring Tool (VEST3) (Carter et al., 2013), Mendelian Clinically Applicable Pathogenicity (M-CAP) (Jagadeesh et al., 2016), deleterious annotation of genetic variants using neural networks (DANN) (Quang et al., 2015), Eigen-PC (Ionita-Laza et al., 2016), Genomic Evolutionary Rate Profiling (GERP++ and GERPP++ GT2) (Davydov et al., 2010), phylogenetic P-values (phyloP100way and phyloP20way) (Pollard et al., 2010), phastCons100way and phastCons20way (Siepel et al., 2005), Site-specific Phylogenetic analysis (SiPhy) (Garber et al., 2009; Lindblad-Toh et al., 2011), REVEL (Ioannidis et al., 2016), and MPC (Samocha et al., 2017).

Method Details

Kinship analysis

Pedigree information and participant relationships were confirmed utilizing pairwise PLINK identity-by-descent (IBD) calculation (Purcell et al., 2007). The IBD sharing between the probands and parents in all trios is between 45% and 55%. Pairwise individual relatedness was corroborated using KING (Manichaikul et al., 2010).

Haplotype phasing and analysis of inbreeding

Haplotype phasing, inbreeding coefficient, and the longest homozygosity-by-descent (HBD) fragment were estimated using Beagle v3.3.2 (Browning and Browning, 2007) as described previously (Jin et al., 2017). The criteria of consanguinity are defined as runs of homozygosity in segments of 2cM or greater length that collectively comprise at least 0.35% of the genome.

Principal component analysis

In order to determine the ethnicity of each participant, we utilized the EIGENSTRAT software (Price et al., 2006) to analyze SNPs in cases, controls, and HapMap subjects as previously described (Jin et al., 2017; Timberlake et al., 2016).

De novo and inherited (dominant/recessive) variant analysis and filtering

De novo mutations were called in parent-offspring trios, each consisting of an affected child with primary congenital hydrocephalus and his or her unaffected biological parents, using the Bayesian framework TrioDeNovo (Dong et al., 2015). Candidate _de novo_ variants were filtered based on the following criteria: (1) minor allele frequency (MAF) $\leq 5 \times 10^{-3}$ in ExAC, 1000 Genomes, and EVS, (2) GATK variant quality score recalibration (VQSR) of 'pass', (3) minimum sequencing depth of 8 reads in the proband and each parent, (4) genotype quality (GQ) score ≥ 20 and alternate allele ratio $\geq 40\%$, (5) TrioDeNovo data quality (DQ) score ≥ 7, and (6) exonic or splice-site variant.

Transmitted dominant variants were filtered by similar criteria of rareness and quality: (1) MAF $\leq 2 \times 10^{-5}$ in ExAC, (2) GQ ≥ 20 and alternate allele ratio $\geq 40\%$, (3) GATK VQSR of 'pass', and (4) minimum sequencing depth of 8 reads in each participant. Recessive variants were also filtered for rare (MAF $\leq 10^{-3}$ in ExAC) bi-allelic events (homozygous and compound heterozygous mutations) that met read quality criteria as above (GQ ≥ 20, alternate allele ratio $\geq 40\%$, 'pass' GATK VQSR, and minimum sequencing depth ≥ 8). Hemizygous recessive variants were filtered for rare events (MAF $\leq 2 \times 10^{-5}$) using the same quality criteria described above.

The impact of nonsynonymous single nucleotide variants on protein function was predicted using the MetaSVM algorithm (Dong et al., 2015), identifying mutations with rank scores greater than 0.83357 as deleterious ('D-mis'). D-Mis and loss-of function mutations (nonsense, frameshift insertions and deletions, and splice-site) were considered potentially damaging to protein function.

All _de novo_ and transmitted calls were verified by _in silico_ visualization of aligned reads using the BLAT search (Kent, 2002) and Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Salient _de novo_ and compound heterozygous calls were then verified in all participants that provided DNA by direct Sanger sequencing of PCR amplicons containing the mutation.

Burden of _de novo_ mutations

The burden of _de novo_ mutations in congenital hydrocephalus cases and unaffected autism controls was determined using the denovolyzeR package (Ware et al., 2015) as previously described The probability of observing a _de novo_ mutation in each gene was calculated as illustrated previously (Jin et al., 2017), with the exception that the coverage adjustment factor was based on the full set of 125 case trios (or 126 case trios in the SMARCC1 analysis given the inclusion of a _de novo_ in parent Hydro106-3) and 1,789 control trios (separate probability tables for each set). The expected number of _de novo_ mutations across variant classes in case and control cohorts was calculated and compared to the observed number of _de novo_ mutations in

each cohort using the Poisson test (Samocha et al., 2014). Gene-set enrichment analyses and statistical tests considered only mutations observed or expected in genes within the specified set (i.e. high brain-expressed, aqueductal stenosis).

Contribution of *de novo* mutation to congenital hydrocephalus

The number of *de novo* mutations in risk genes that contribute to congenital hydrocephalus was calculated based on the observed count of protein-altering *de novo* mutations compared to expectation (=Nx((M_1-M_2)/M_1), where N is the total number of trios and M_1 and M_2 are the observed and expected count of protein-altering *de novo* mutations per trio, respectively)

Enrichment analysis for the dominant and recessive variants

To quantify the enrichment of LoF heterozygous variants, we calculated the expectation for a gene using the following formula:

$$\text{Expected LoF}_j = L \times \frac{\text{mutability}_j}{\sum_{Genes} \text{mutability}_j}$$

where 'j' denotes the 'jth' gene and 'L' denotes the total number of LoF heterozygous mutations. A one-tailed binomial test was conducted to compare the observed number of heterozygous variants to expectation.

For damaging recessive genotypes (RGs) in a specific gene in cases, we conducted a one-tailed binomial test to evaluate enrichment as described previously (Jin et al., 2017). RG can also be modeled separately as compound heterozygotes or homozygotes. The expected number of compound heterozygotes for each gene is derived from distributing the observed number of RGs, N, across all genes according to the ratio of the squared *de novo* probabilities:

$$\text{Expected Compound RG}_i = N \times \frac{probability^2_{de\ novo}}{\sum_{Genes}(probability^2_{de\ novo})}$$

The expected number of homozygotes is derived similarly, but using the linear ratio of *de novo* probabilities:

$$\text{Expected Homozygous RG}_i = N \times \frac{probability_{de\ novo}}{\sum_{Genes}(probability_{de\ novo})}$$

The total number of expected RG for each gene is the sum of the derived expected compound heterozygous and homozygous values.

Gene ontology enrichment analysis.

Three genome-wide significant genes *PTCH1*, *SMARCC1*, and *PTCH1* were input into GOrilla (Eden et al., 2009) to identify enriched GO terms compared to the background set of genes

(M=18,715). For gene-set enrichment analyses, each statistical test considered observed or expected mutations in genes within the specified gene set.

In Silico Splice-Site Prediction

In order to assess the impact of a missense mutation at the splice donor site of intron 10 of PTCH1 that changed the donor site sequenced from GTA to GTC, we utilized Human Splicing Finder (Desmet et al., 2009) and MaxEntScan (Yeo and Burge, 2004). Both programs predicted that this mutation is likely to affect splicing. MaxEntScan assigned a MaxEnt score of 7.64 to the wild-type canonical splice donor site; however, after mutation of the wild-type sequence to GTC, the alternate splice donor bases upstream was assigned a score of -0.18. The difference in MaxEnt scores between the mutant and the reference sequence is -7.82, which provides strong support for this mutation being a potential 5' donor splice site.

Copy Number Variation (CNV) analysis

XHMM was run to call CNVs from WES as previously described (Ruderfer et al., 2016). GATK DepthOfCoverage was used to calculate mean read depth per targets from the alignment files. The data was normalized by removing the highest variance principal components (variance > 70%) and z scores were calculated from the mean read depths. CNVs were called using the Viterbi Hidden Markov model (HMM) and the quality scores were calculated using the forward-backward HMM. After filtering out common CNVs present at allele frequencies greater than 0.1% in 1000 Genomes (Sudmant et al., 2015) and 10% in the cohort, high quality CNVs (SQ>60 where SQ indicates the phred-scaled quality score for the presence of a CNV event within the interval) were subjected to visual inspection. *De novo* CNVs were assessed using PLINK/Seq command-line tools (Fromer et al., 2012).

Quantification and Statistical Analyses

De novo CNV Permutation Test

The probability of finding one exon covered by multiple *de novo* CNVs was calculated by comparing the observed distribution to an empirical distribution derived from 1 million permutations. For each permutation, five *de novo* duplications were randomly distributed across the genome; each duplication contained the same number of exons as predicted from XHMM with adjustment if the CNV is partially validated by experiment. In each permutation, the experiment was considered a success if at least one locus contained exactly one exon covered by multiple duplications. The number of successes was tallied and the p-value was obtained by dividing the number of successes by the number of iterations.

Estimating the number of risk genes

A maximum likelihood approach was used to estimate the number of genes contributing to congenital hydrocephalus via de novo events as described previously (Jin et al., 2017). We defined K to be the number of observed protein-altering de novo mutations in high brain-expressed (HBE) genes among cases. R1 indicates the number of HBE genes mutated exactly

twice in cases and R2 indicates the number of high HBE mutated three times or more. We set the proportion (E) of protein-altering mutations in risk genes based on point estimate of enrichment in cases compared to expectation (E = (M1-M2)/M1, where M1 and M2 are the observed and expected count of protein-altering de novo mutations per trio, respectively). We then simulated the likelihood function as follows: First, we randomly selected G risk genes from the HBE gene set. Next, we simulated the number of contributing protein-altering mutations in risk genes, i.e. C, by sampling once from Binomial(K,E) distribution. Then, we simulated C contributing protein-altering mutations in G risk genes and K-C non-contributing protein-altering mutations in the complete HBE gene set, using each gene's protein-altering mutability score as probability weights. We performed 20,000 simulations for G from 2 to 100, and calculated the likelihood function L(G) as the proportion of simulations in which the number of genes with two protein-altering de novo mutations equals to R1 and the number of genes with three or more protein-altering de novo mutations equals to R2. We then estimated the number of risk genes using the maximum likelihood estimate (MLE). Based on the likelihood function, we calculated the Fisher information and constructed the confidence interval based on the MLE and estimated Fisher information using the following equation

$$\text{MLE} \pm 1.96 \times \left(\frac{1}{\sqrt{\text{Fisher Information}}} \right)$$

In situ hybridization

Mouse brains and embryos were fixed in 4% paraformaldehyde by overnight immersion and sectioned (10-15 μm, cryostat sections for digoxigenin [DIG] probes). Antisense RNA probes corresponding to murine Ptch1, Smarcc1 and Trim71 (approx. 200bp for DIG-labeled probes respectively) were synthesized to detect Ptch1, Smarcc1 and Trim71 transcripts in murine tissue, using methods previously described (Duncan et al., 1994; Petersen et al., 2002; Schaeren-Wiemers and Gerfin-Moser, 1993).

Data and Software Availability

The sequencing data for congenital hydrocephalus case-parent trios reported in this manuscript have been deposited in the NCBI database of Genotypes and Phenotypes (dbGaP) (accession no. phs000744).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Whole exome sequencing data from CH trios (n = 125)	This paper	Accession no. phs000744 (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000744)
Whole exome sequencing data from SSC control trios	Iossifov et al., 2014	NDAR: DOI:10.15154/1149697 (https://ndar.nih.gov/study.html?id=352)
Software and Algorithms		
Genome Analysis Tool Kit (GATK)	DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2013	https://software.broadinstitute.org/gatk/best-practices/
BWA-mem	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
Annovar	Wang et al., 2010	http://annovar.openbioinformatics.org/en/latest/
PLINK/SEQ	Fromer et al., 2014	https://atgu.mgh.harvard.edu/plinkseq/
Other		
1000 Genomes GRCh37 h19 genome build	1000 Genomes Project	http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/human_g1k_v37.fasta.gz
RefSeq hg19 gene annotation	UCSC Genome Browser	http://genome.ucsc.edu/cgi-bin/hgTables?command=start
Intervals file for IDT xGen v1.0	Integrated DNA Technologies	https://www.idtdna.com/pages/products/next-generation-sequencing/hybridization-capture/lockdown-panels/xgen-exome-research-panel

Main Figures

Figure 1. Recurrent, identical *de novo* mutations in *LIN41TRIM71*, encoding the let-7 miRNA target TRIM71.

a. Representative sagittal (left) and axial (right) T2-weighted brain magnetic resonance images of CH probands Hydro100-1 and Hydro102-5 showing severe communicating hydrocephalus. Hydro100-1 and Hydro102-5 were both treated with surgical CSF shunting.

b. Pedigree structures with Sanger-verified mutated bases (red) and the corresponding wild-type bases marked on the chromatograms.

c. Structural modeling of TRIM71 mutation impact on surface electrostatic charge patterns in the protein. Normally, the positively charged guanidinium side chain of p.Arg796 interacts with the negatively charged sugar-phosphate backbone of target RNAs to aid in maintaining the spatial position of the nucleic acid. Mutation of this p.Arg796 residue to the imidazole ring of histidine ($\Delta\Delta G$ = 2.0 Kcal/mol) is predicted to disrupt this these interactions (right). The side chain of p.Arg608 makes hydrogen bonds with uracil in target RNAs. The p.Arg608His mutation ($\Delta\Delta G$ = 1.6 Kcal/mol) is predicted to result in the loss of these hydrogen bonds (left).

d. Locations of identified TRIM71 mutations in relation to critical functional domains. p.Arg608His and p.Arg796His both affect conserved residues in the 16th position of the respective 1st and 5th blades of TRIM71's NHL domain, which mediates the binding of target RNA.

e. *In situ* hybridization of wild-type E12.5 and adult mouse brains for *Trim71* showing signals in the ciliated neuroepithelium lining the developing neural tube and ventricular zone (V, ventricle; arrow, neuroepithelium at E12.5 and ependymal layer in adulthood; 2.5x and 10x magnification).

Figure 1. Recurrent, identical *de novo* mutations in *LIN41/TRIM71*, encoding the let-7 miRNA target TRIM71.
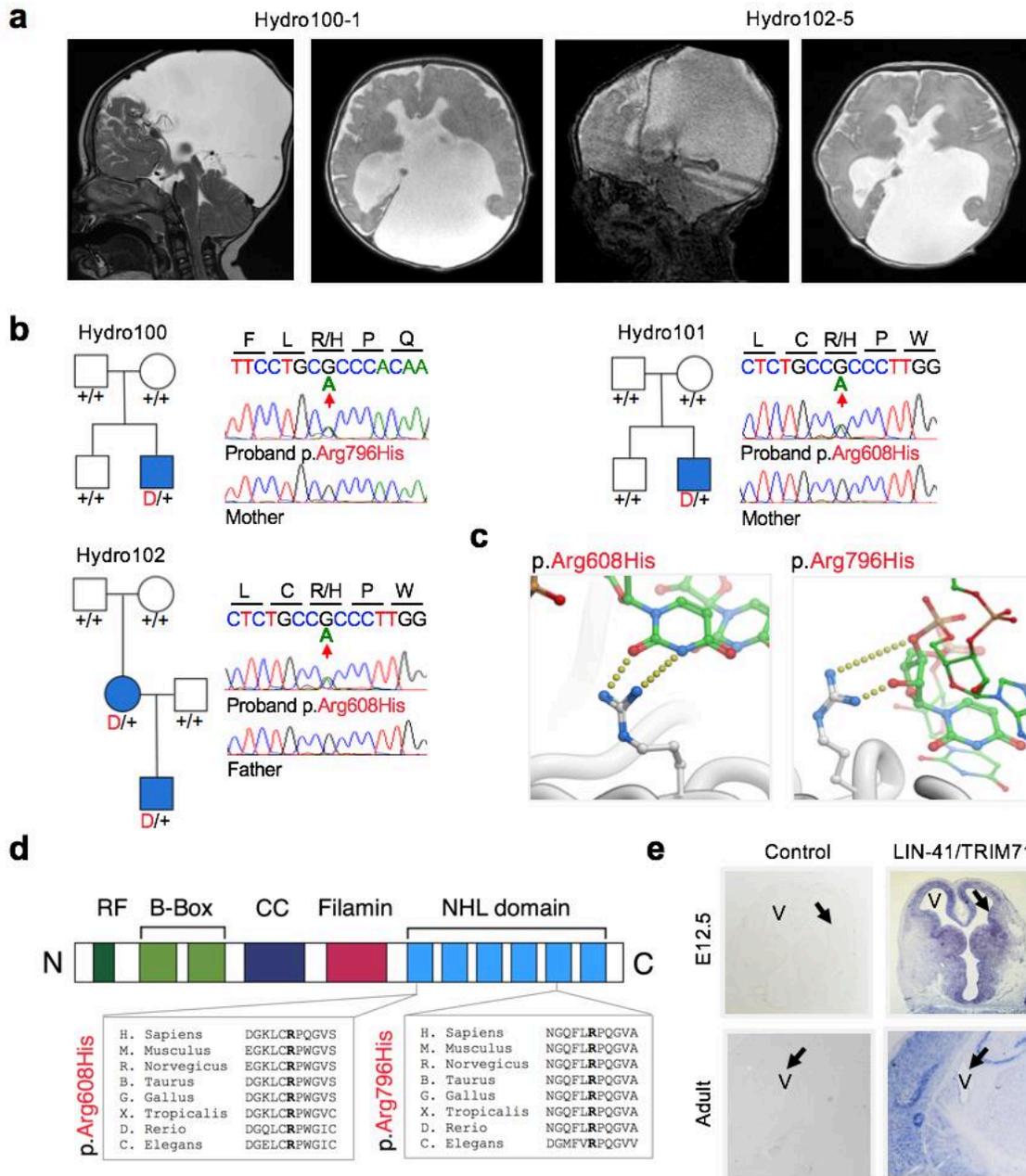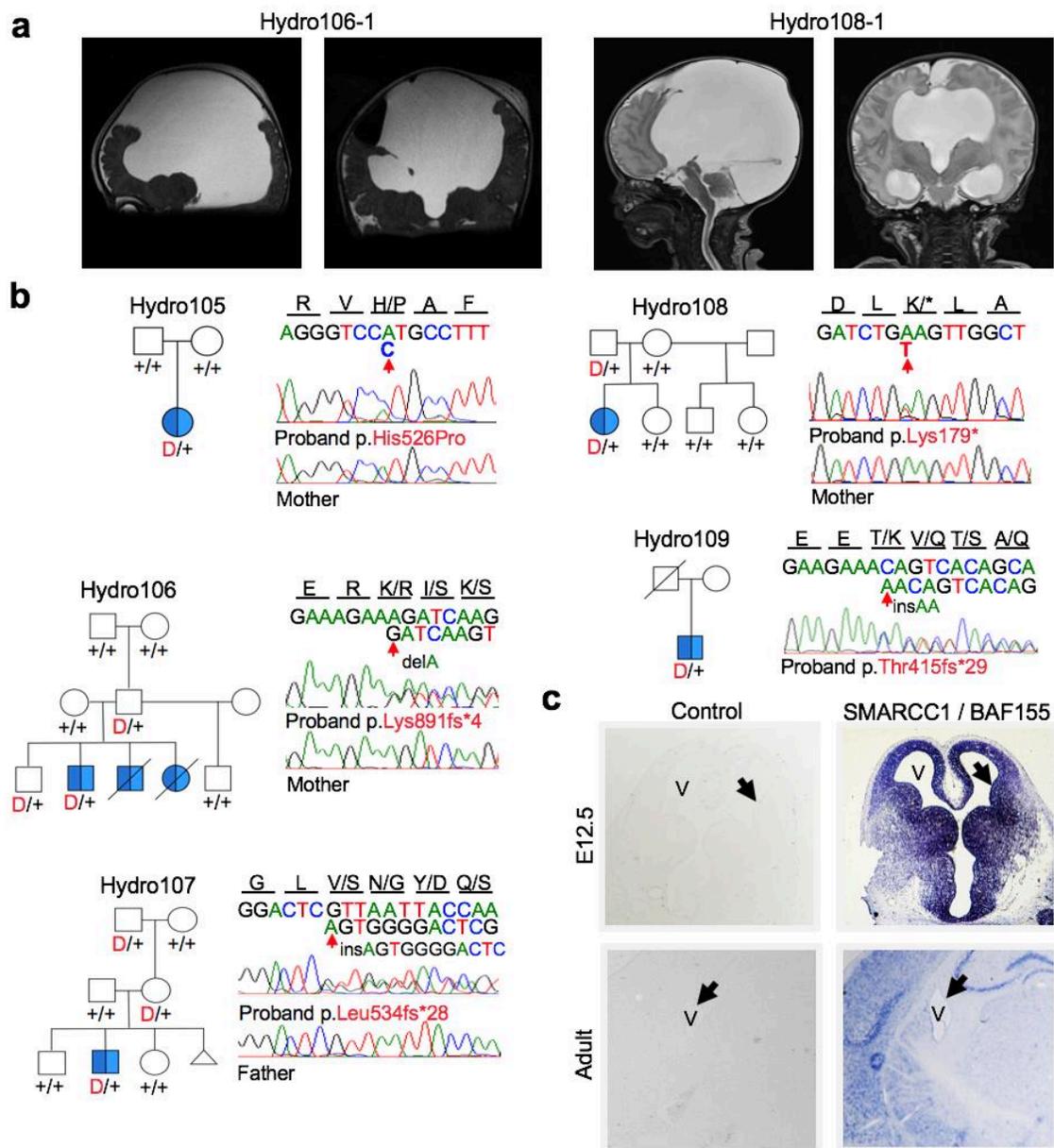
Figure 2. Multiple *de novo* and transmitted mutations in *SMARCC1*, encoding the SWI/SNF chromatin modifier, SMARCC1 (BAF155).

a. Representative sagittal (left) and coronal (right) T2-weighted brain magnetic resonance images of CH probands Hydro106-1 and Hydro108-1 demonstrate severe obstructive hydrocephalus. All three probands required surgical CSF diversion.
b. Pedigree structures with Sanger-verified mutated bases and the corresponding normal alleles marked on the chromatograms.
c. *In situ* hybridization of wild-type E12.5 and adult mouse brains for *Smarcc1* showing signals in the ciliated neuroepithelium lining the neural tube and ventricular zone (V, ventricle; arrow, neuroepithelium at E12.5 (2.5x) and ependymal layer in adulthood; (10x magnification).

Figure 3. Multiple *de novo* and transmitted mutations in *PTCH1*, encoding the Sonic Hedgehog receptor Patched-1.

a. Representative sagittal (left) and axial (right) T2-weighted brain magnetic resonance images of CH probands Hydro103-1 and Hydro104-1 with obstructive hydrocephalus, both treated with surgical shunting.
b. Pedigree structures with Sanger-verified mutated bases and the corresponding normal alleles marked on the chromatograms.
c. *In situ* hybridization staining of wild-type E12.5 and adult mouse brains for *Ptch1* demonstrates specific expression in the hindbrain ciliated neuroepithelium (V, ventricle; arrow, neuroepithelium at E12.5 (2.5x) and ependymal layer in adulthood (10x magnification).
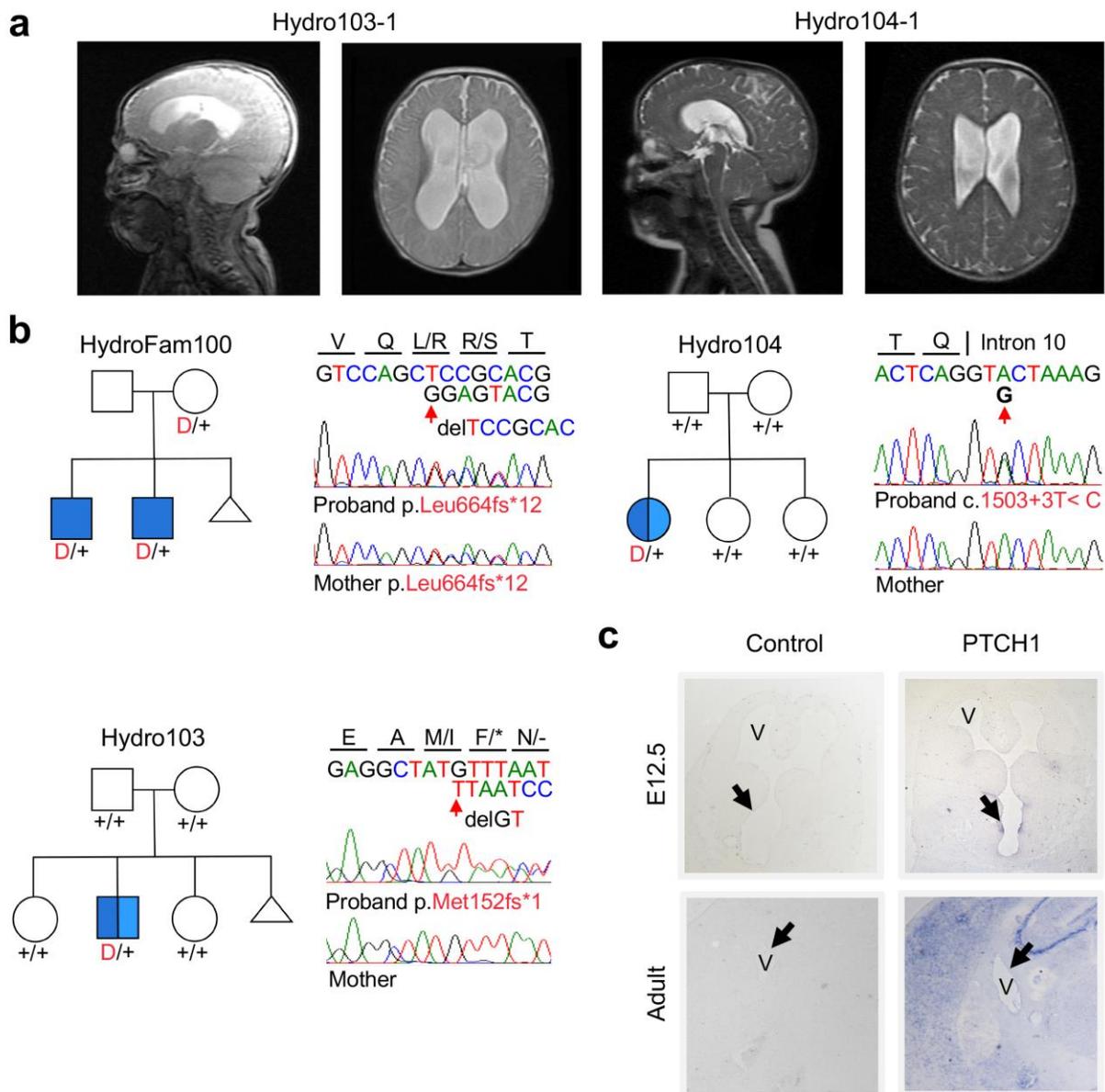
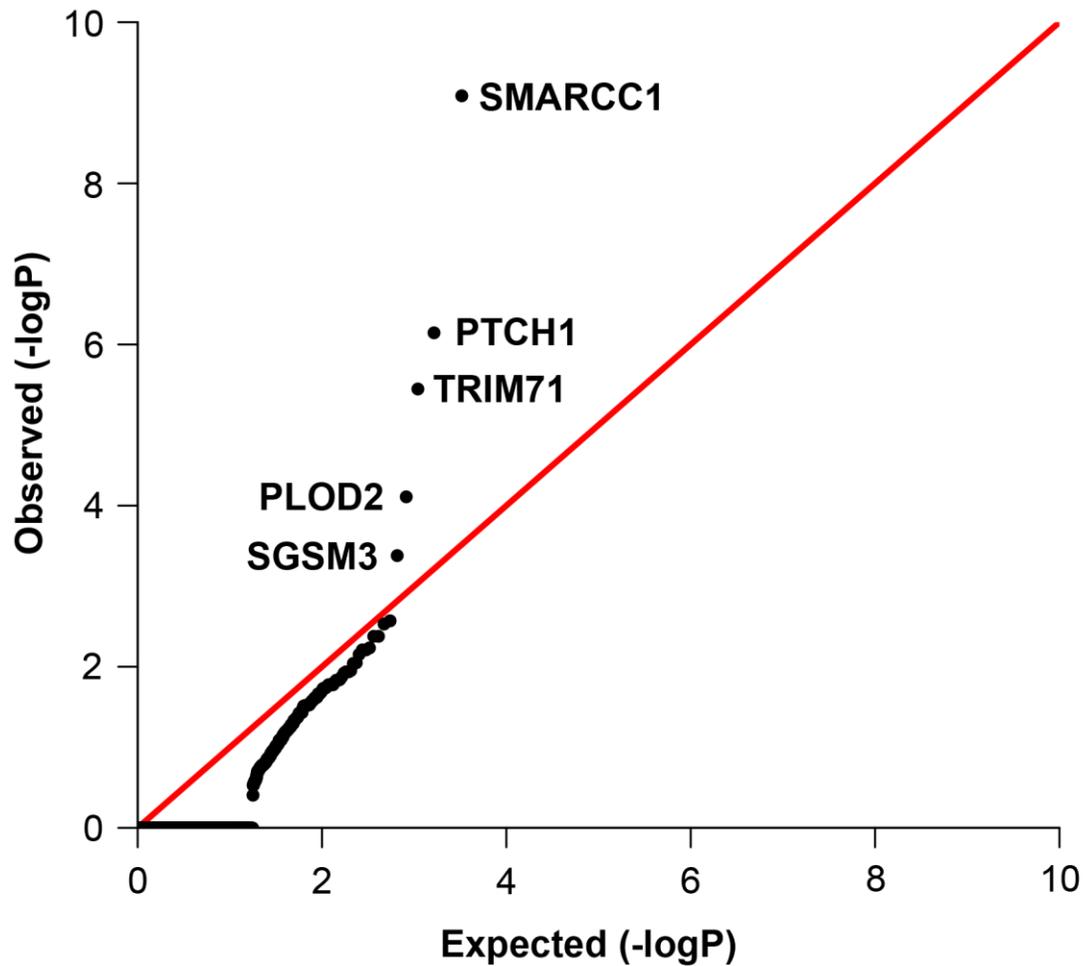**Figure 4. Meta-analysis of rare *de novo* and transmitted mutations in mutation-intolerant genes.** Quantile-Quantile plot of observed versus expected p-values from meta-analysis of protein-altering *de novo* and loss-of-function transmitted heterozygous variants, comparing the burden of rare variants in genes intolerant to mutation (pLI > 0.9, MAF < 2 x 10^{-5})

Main Tables

Table 1. Observed and expected *de novo* mutation rates in 125 probands with congenital hydrocephalus (CH) and CH cases with aqueductal stenosis (AS) for variant classes by phenotype and high brain expression.

	Observed		Expected		Enrichment	p
	n	Rate	n	Rate		
Hydro (125)						
All	137	1.10	139.7	1.12	0.98	0.60
Synonymous	32	0.26	39.7	0.32	0.81	0.91
Missense	89	0.71	87.8	0.70	1.01	0.46
D-Mis	16	0.13	16.5	0.13	0.97	0.58
LOF	16	0.13	12.3	0.10	1.30	0.18
Protein-Altering	105	0.84	100.0	0.80	1.05	0.32
Damaging	32	0.26	28.7	0.23	1.11	0.30
Hydro (125) - High Brain Expressed						
All	48	0.38	36.9	0.30	1.30	**0.04***
Synonymous	12	0.10	10.3	0.08	1.16	0.34
Missense	30	0.24	23.2	0.19	1.29	0.10
D-Mis	5	0.04	4.4	0.04	1.14	0.45
LOF	6	0.05	3.4	0.03	1.78	0.13
Protein-Altering	36	0.29	26.5	0.21	1.36	**0.04***
Damaging	11	0.09	7.8	0.06	1.42	0.16
Hydro + AS (66)						
All	78	1.18	73.8	1.12	1.06	0.33
Synonymous	22	0.33	20.9	0.32	1.05	0.44
Missense	47	0.71	46.3	0.70	1.01	0.48
D-Mis	12	0.18	8.7	0.13	1.38	0.17
LOF	9	0.14	6.5	0.10	1.39	0.21
Protein-Altering	56	0.85	52.8	0.80	1.06	0.35
Damaging	21	0.32	15.2	0.23	1.38	0.09
Hydro + AS (66) - High Brain Expressed						
All	30	0.45	19.5	0.30	1.54	**0.01***
Synonymous	9	0.14	5.5	0.08	1.65	0.10
Missense	17	0.26	12.2	0.19	1.39	0.11
D-Mis	4	0.06	2.3	0.04	1.72	0.21
LOF	4	0.06	1.8	0.03	2.25	0.11
Protein-Altering	21	0.32	14.0	0.21	1.50	**0.04***
Damaging	8	0.12	4.1	0.06	1.95	**0.05***

n: the number of *de novo* mutations; Rate: the number of *de novo* mutations per individual; Enrichment: ratio of observed to expected number of mutations; High Brain Expression: top quartile of expression; Missense: tolerated missense mutations as predicted by MetaSVM; D-Mis: damaging missense mutations as predicted by MetaSVM; LOF: loss-of-function denotes premature termination, frameshift, or splice site mutations; Damaging: D-miss and LOF mutations

CellPress

Table 2. Meta-analysis of protein-altering *de novo* mutations (DNMs) and loss-of-function (LOF) heterozygous mutations in probands for genes with multiple DNMs.

Gene	*De Novo* Mutations				Transmitted Loss-of-Function (LOF)				Meta P-value	pLI	High Brain Expression Rank
	# observed	# expected	enrichment	p-value	# observed	# expected	enrichment	binomial p-value			
SMARCC1	2	2.3×10^{-3}	861.7	**2.7×10^{-6}**	3	4.3×10^{-2}	70.5	1.2×10^{-5}	**8.2×10^{-10}**	1.0	97.4
PTCH1	2	1.5×10^{-3}	1374.3	**1.1×10^{-6}**	1	3.8×10^{-2}	26.1	3.8×10^{-2}	**7.2×10^{-7}**	1.0	93.2
TRIM71	3	1.1×10^{-2}	274.8	**2.2×10^{-7}**	0	2.5×10^{-2}	0	1	3.5×10^{-6}	1.0	74.6
PLOD2	2	3.5×10^{-3}	580.2	5.9×10^{-6}	0	7.8×10^{-2}	0	1	7.7×10^{-5}	0	82.8
SGSM3	2	8.6×10^{-3}	232.2	3.7×10^{-5}	0	6.9×10^{-2}	0	1	4.1×10^{-4}	0	67.4

Meta-analysis performed by combining the p-values from protein-altering *de novo* mutations and loss-of-function (LOF) heterozygous mutations using the Fisher's method with 4 degrees of freedom. Bolded p-values surpass the Bonferroni multiple testing correction (2.6×10^{-6}, 0.05/18,989) for p-values tabulated by either *de novo*, heterozygous or meta-analysis. Of note, the denovolyzeR p-value of damaging SMARCC1 mutations was calculated using 126 case-parent trios given the sporadic *de novo* mutation occurs in the proband's father (Hydro106-3).

43

Supplementary Figures

Supplementary figure 1. *De novo* **mutation rates closely approximate a Poisson distribution in cases and controls.**

Observed number of *de novo* mutations per case-parent trio (bars) compared to expected (line) from the Poisson distribution in CH case (red) and healthy autism sibling control (blue) cohorts; 'p' denotes the chi-squared p-value.
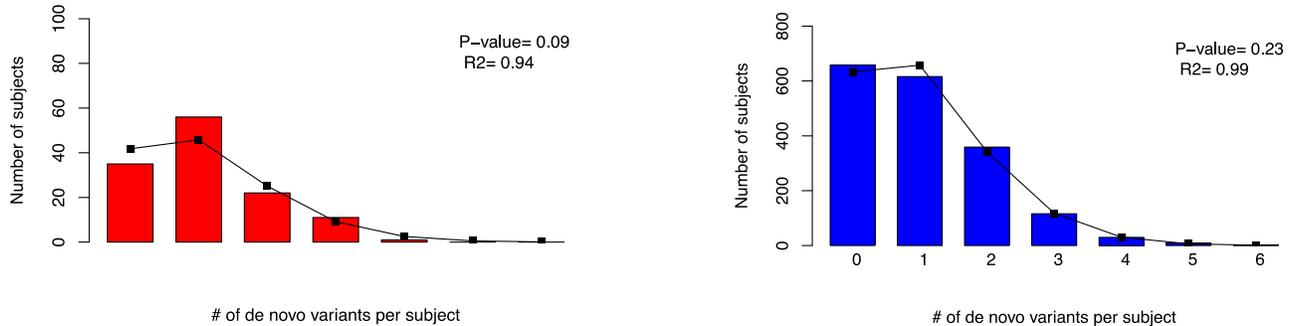
Supplemental Figure 2. Principal-component analysis of our hydrocephalus cohort of 440 individuals (+) clustered along with HapMap subjects (o).
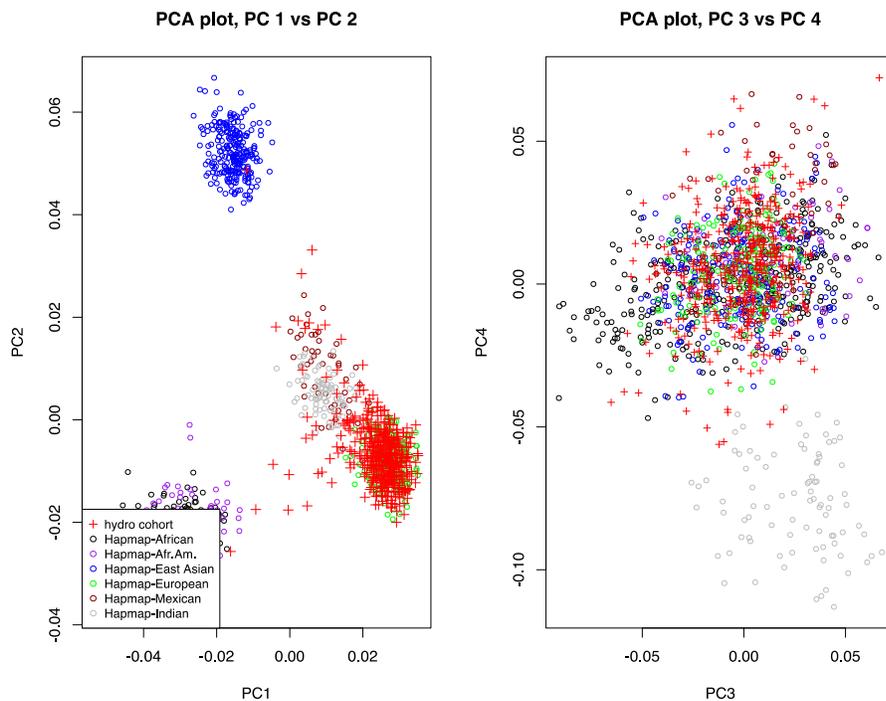
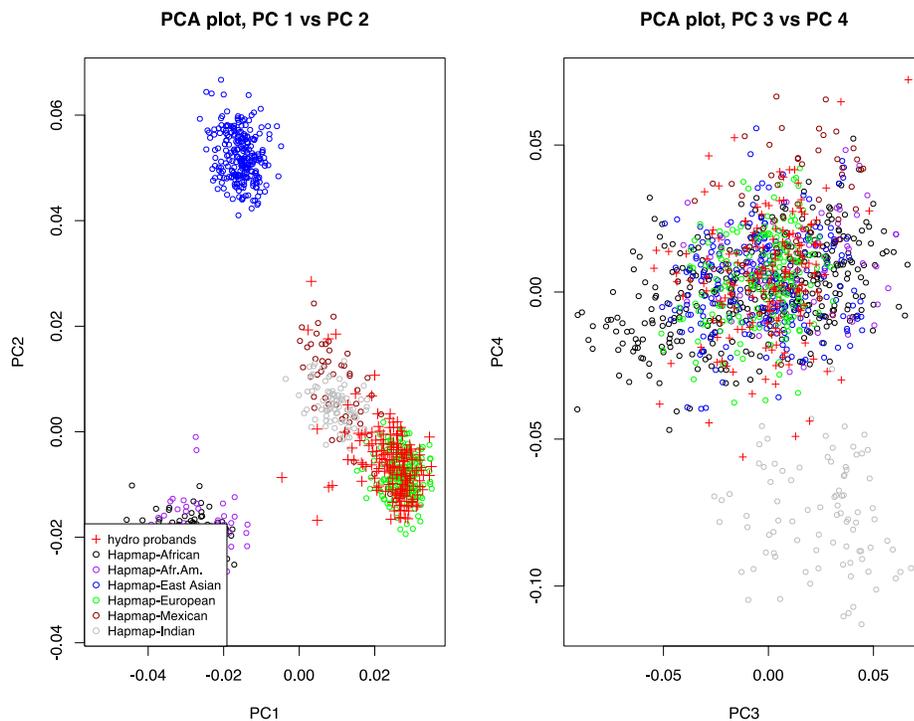**Supplemental Figure 3. Principal-component analysis of 177 probands with congenital hydrocephalus clustered along with HapMap subjects**. Results identify 162 probands (+) that cluster with HapMap Non-Finnish, European subjects (o).

Supplementary Figure 4. (A) Homology model of TRIM71 (residues 577-868) based on the crystal structure of the Brat-NHL domain. TRIM71 belongs to the NHL superfamily exhibiting structural conservation of the hexa-NHL repeat. R608 and R796 are shown as cyan sticks. (B) Surface representation of TRIM71 (white surface with R608 and R796 in blue) with bound 14-nucleotide RNA (orange backbone with nucleotides as green sticks).

45
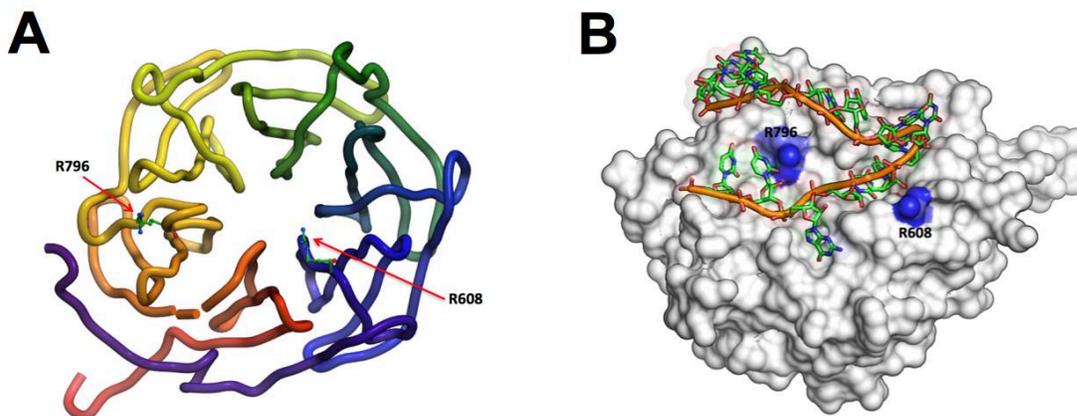
Supplementary Figure 5. (A) The positively charged TRIM71 guanidinium side chain of p.Arg796 interacts with the sugar-phosphate backbone of its target RNA, possibly stabilizing the spatial orientation of the nucleic acid. The imidazole ring of the missense mutant histidine ($\Delta\Delta G$ = 2.0 Kcal/mol) is unable to retain these interactions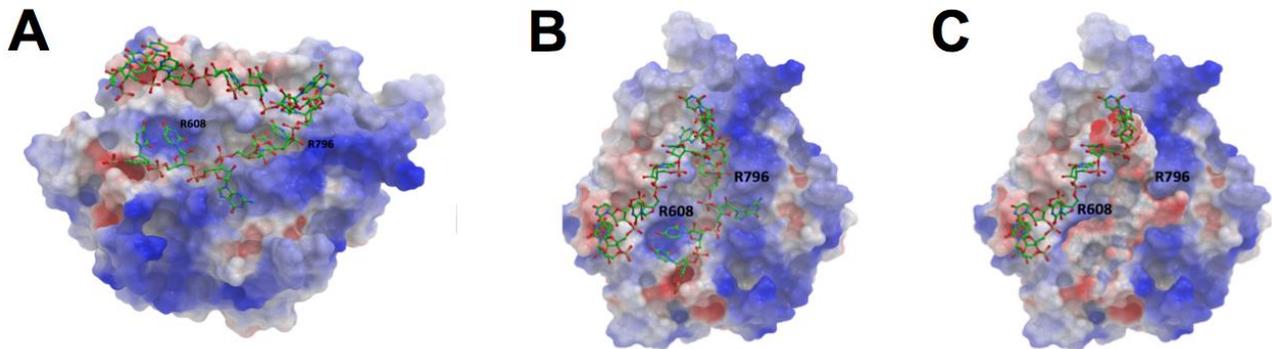. (B) The R608 guanidinium side chain makes hydrogen bonds with a uracil base of its target RNA. The p.Arg608His mutation ($\Delta\Delta G$ = 1.6 Kcal/mol) results in the loss of these hydrogen bonds. (C) TRIM71 (cyan surface) and TRIM71 residues (yellow surface) within 5Å of bound RNA (cpk sticks) with spatial positions of p.Arg608 and p.Arg796 labeled.
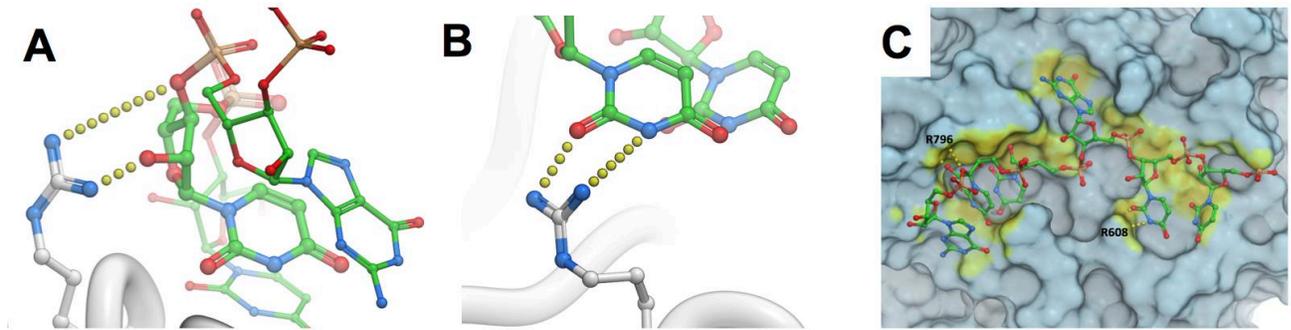
Supplementary Figure 6. (A) Lateral and (B) Superior view of RNA (green) interacting with TRIM71. The surface of TRIM71 is colored based on an electrostatic surface representation. The part of the RNA interacting interface is positively charged (blue). (C) Eight nucleotides that interact with TRIM71 have been illustrated by surface representation and the color highlights the surface charge pattern. It is worth noting the complementary electrostatic charge pattern at the interface of RNA (red) and TRIM71 (blue).

Supplementary Figure 7. (A) Sequence alignment of TRIM71 NHL domain (Uniprot ID Q2Q1W2 residues 577-868) with the sequences from the NHL domain of Brat, which were used as templates for homology modelling (PDB id 1Q7F, 5EX7 and 4ZLR). (B) Overall model quality as determined by Protein structure analysis (PROSA) webserver (Sippl, 1993; Wiederstein and Sippl, 2007). The Z-score is -6.07. The z-score indicates overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. Z-scores outside a range characteristic for native proteins indicate erroneous structures. The black dot represents the final chosen model. (C) Conserved domain architecture of human TRIM71 as predicted by blastp suite at the NCBI server (Altschul et al., 1990; Altschul et al., 1997; Boratyn et al., 2012; Camacho et al., 2009; Madden et al., 1996; Morgulis et al., 2008; States and Gish, 1994; Zhang and Madden, 1997; Zhang et al., 2000). Residues 577-868 in query sequence (Uniprot Id Q2Q1W2) were predicted to be RING-HC_TRIM71_C-VII and NHL_TRIM71_like domain-containing protein (domain architecture ID 11614574) belonging to the NHL superfamily.

Supplementary Figure 8. *In situ* hybridization staining of wild-type mouse embryos for *Trim71*, *Smarcc1*, and *Ptch1* in the embryonic day 12.5 (E12.5, 2.5x) and adult mouse brain (10x magnification) demonstrates specific expression in the ciliated neuroepithelium lining the neural tube and ventricular zone in the developing brain.
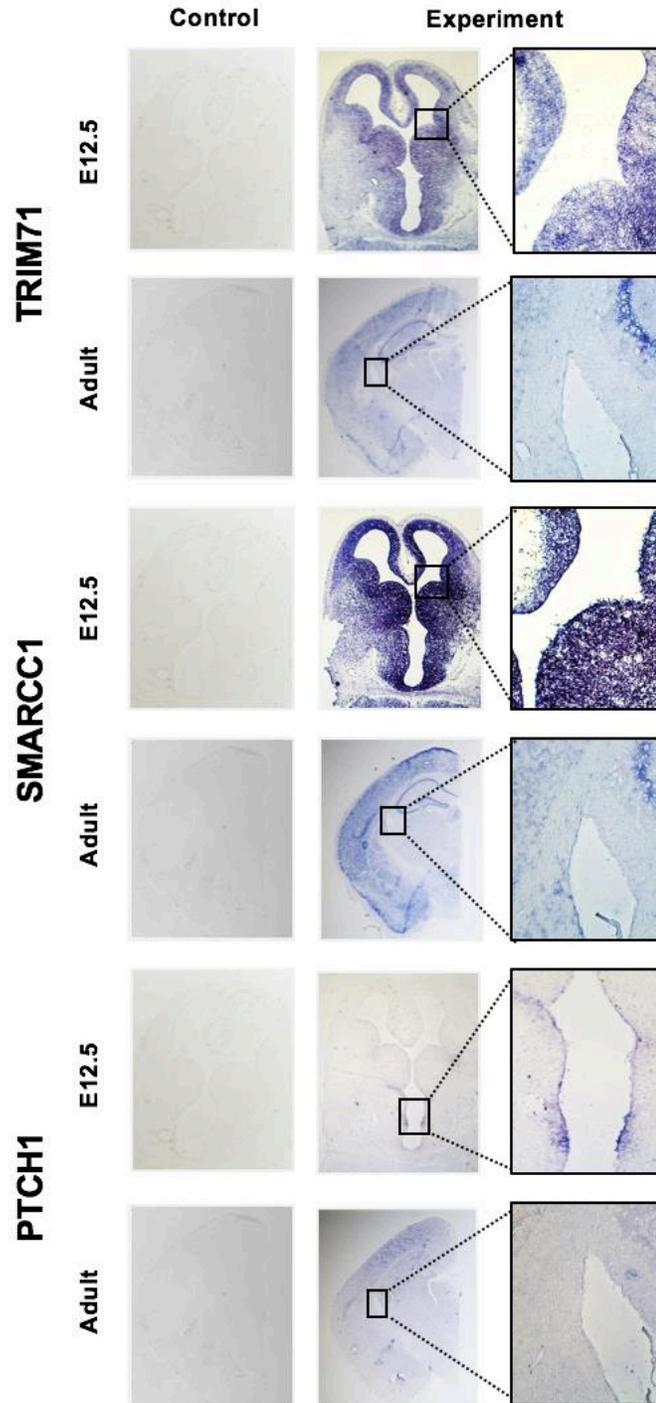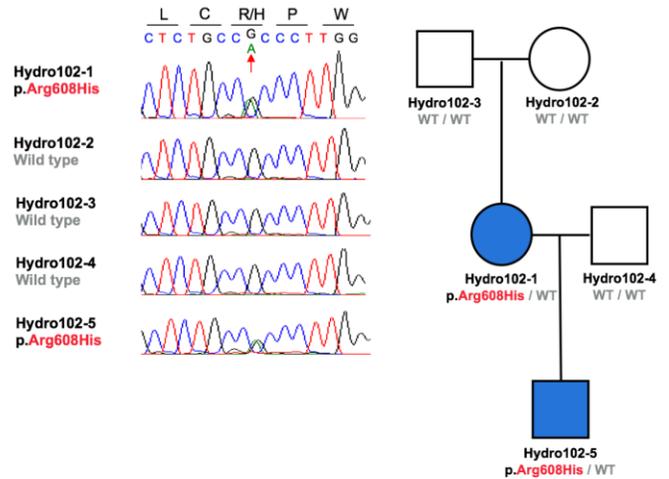
Supplementary Figure 9. Pedigrees and Sanger sequence chromatograms for three families carrying *de novo* mutations in *LIN41/TRIM71*.
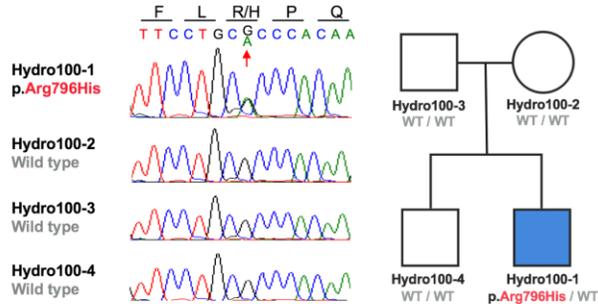
Supplementary Figure 10. Pedigrees and Sanger sequence chromatograms for three families carrying loss-of-function mutations in *PTCH1*.

a. Hydro103: PTCH1 c.456_457delGT: p.Met152fs*1
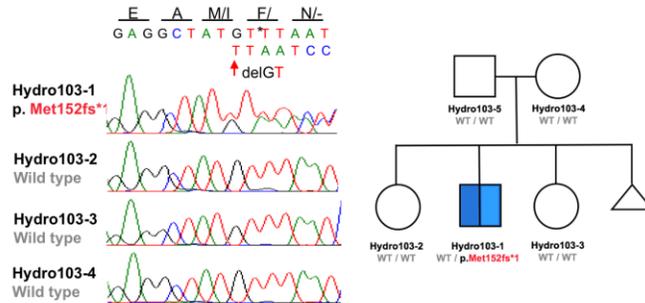
b. Hydro104: PTCH1 c.1503+3T<C

c. HydroFam100: PTCH1 c.1991_1997delTCCGCAC: p.Leu664fs*12
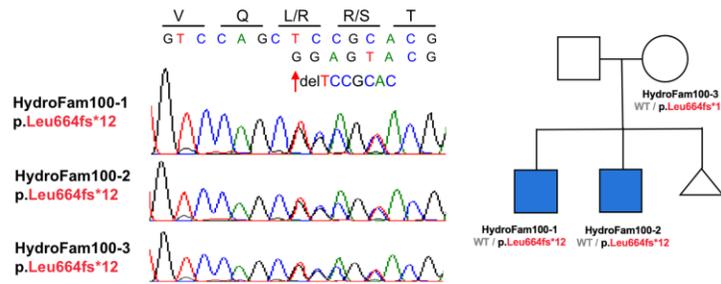
Supplementary Figure 11. Pedigrees and Sanger sequence chromatograms for three families carrying damaging mutations in *SMARCC1*.
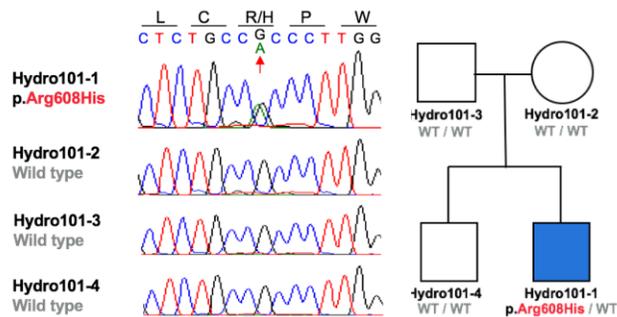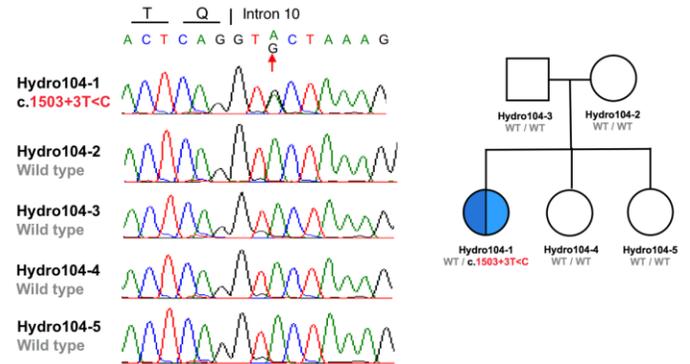
a. Hydro105: SMARCC1 c.1577A>C: p.His526Pro

d. Hydro107: SMARCC1 c.1602_1603insAGTGGGGACTC: p.Leu534fs*2

b. Hydro106: SMARCC1 c.2672delA: p.Lys891fs*4
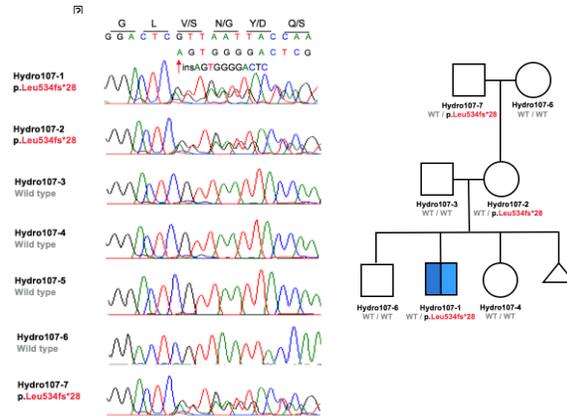
e. Hydro108: SMARCC1 c.535A>T : p.Lys179*

c. Hydro109: SMARCC1 c.1243_1244insAA: p.T415fs*29
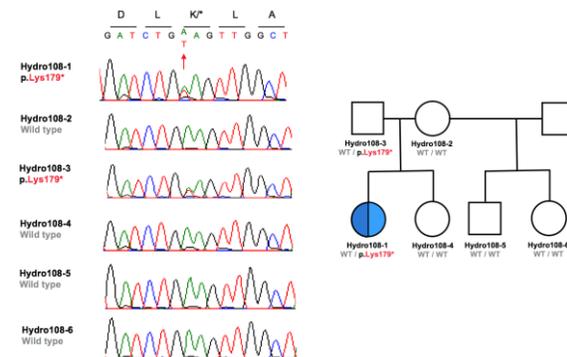
Supplementary Figure 13. Estimation of the number of genes contributing to congenital hydrocephalus via *de novo* mutations. Monte Carlo simulation was performed based on observed protein-altering *de novo* mutations in genes highly expressed in the developing mouse brain using 20,000 iterations. This likelihood-based approach assumes identical penetrance of all protein-altering *de novo* mutations in all risk genes (for details, see Methods). We estimate that the number of risk genes via *de novo* events is ~8 (95% confidence interval: [2.52, 13.48]).
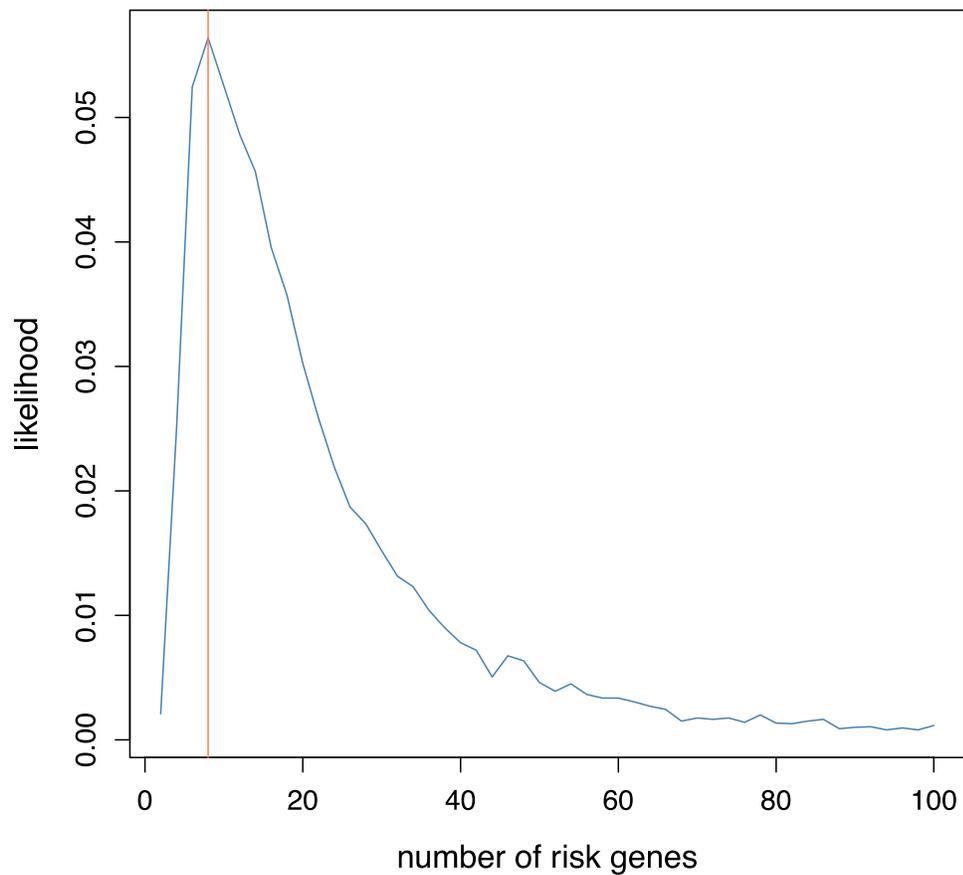
Supplementary Figure 14. Saturation analysis for detecting risk genes with more than one protein-altering *de novo* mutation. The number of trios (x-axis) and the number of risk genes with more than one protein-altering *de novo* mutation (y-axis) are specified. The expected number of risk genes was determined by the maximum likelihood approach to be 8. In each iteration of the simulation, the total number of protein-altering mutations in risk genes was determined based on the point estimate of enrichment in cases compared to expectation (= $N \times [(M1-M2)/M1]$, where N is the number of trios and M_1 and M_2 are the observed and expected count of protein-altering *de novo* mutations per trio, respectively). We then randomly distributed this amount across eight randomly-selected risk genes and calculated how many of these eight genes have > 1 protein-altering *de novo* events. 10,000 iterations were performed to estimate the number of risk genes with more than 1 protein-altering *de novo* mutation, taking into account the protein-altering *de novo* mutation probability. Whole exome sequencing of 500 and 1,000 trios will yield 6.4 and 7.6 risk genes, reaching respective saturations for all risk genes of 80% and 95%.
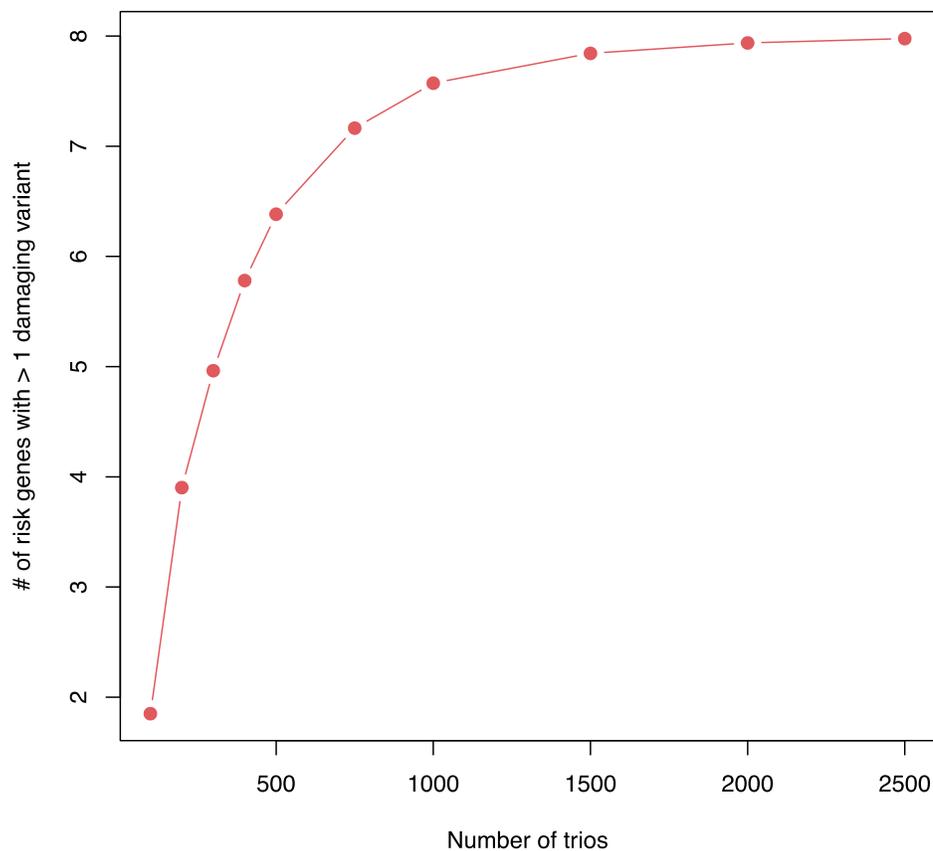
Supplementary Tables

Supplementary Table 1. Number of studied congenital hydrocephalus cases and controls

Category	N
# of unrelated congenital hydrocephalus probands	177
# of case-parent trios	125
# of singleton cases	47
# of familial forms	5
# of control case-parent trios	1,789

N: Number of case-parent trios, singleton cases, and familial forms (defined as greater than one patient in nuclear family affected by congenital, primary hydrocephalus)

Supplementary Table 2. Demographic characteristics of congenital hydrocephalus (CH) cases and controls

	CH Cases	Autism Cohort Sibling Controls
Sample Size	177	1,789
Gender		
Male	90 (50.8%)	840 (47.0%)
Female	87 (49.2%)	949 (53.0%)
Ethnicity		
European	162 (91.5%)	1,408 (78.7%)
Mexican	11 (6.2%)	79 (4.4%)
African American	1 (0.56%)	41 (2.3%)
East Asian	1 (0.56%)	87 (4.9%)
South Asian	2 (1.0%)	125 (7.0%)
Other	0 (0.0%)	49 (2.7%)

The number of samples is shown in each category with the corresponding percentage in parenthesis. Ethnicity is determined by principal component analysis compared to HapMap samples using EIGENSTRAT (see **Methods**).

Supplementary Table 3. Summary sequencing statistics for the congenital hydrocephalus (CH) and control cohorts

Category	CH Cohort (xGen IDT; N = 440)	Control Cohort (Roche V2; N = 5,367)
Read length (bp)	99	50-94
# of reads per sample (M)	50.99	99.7
Median coverage at each targeted base (X)	49.91	67
Mean coverage at each targeted base (X)	54.33	79.1
% of all reads that map to target	57.64%	57.6%
% of all bases that map to target	43.81%	49.0%
% of targeted bases read at least 8x	95.90%	94.6%
% of targeted bases read at least 10x	95.28%	93.4%
% of targeted bases reads at least 15x	92.70%	89.9%
% Mean error rate	0.29%	0.4%

All hydrocephalus cohort samples were sequenced using the xGen IDT capture reagent. 8X, 10X and 15X target base coverage was comparable across platforms.

Supplementary Table 4. Observed and expected *de novo* mutation rates in 1,789 healthy autism controls for variant classes

	Observed		Expected		Enrichment	p
	n	Rate	n	Rate		
Control, Healthy Autism Sibling Cohort						
All	1830	1.02	1949.9	1.09	0.94	1.00
Synonymous	484	0.27	549.6	0.31	0.88	1.00
Missense	974	0.54	993.3	0.56	0.98	0.81
D-Mis	222	0.12	232.8	0.13	0.95	0.77
LOF	150	0.08	174.3	0.10	0.86	0.97
Protein-Altering	1346	0.75	1400.3	0.78	0.96	0.93
Damaging	372	0.21	407.1	0.23	0.91	0.96

n: the number of *de novo* mutations; Rate: the number of *de novo* mutations per individual; Enrichment: ratio of observed to expected numbers of mutations; Missense: tolerated missense mutations as predicted by MetaSVM; D-Mis: damaging missense mutations as predicted by MetaSVM; LOF: loss-of-function denotes premature termination, frameshift or splice site mutation; Damaging: D-Mis and loss-of-function mutations

Supplementary Table 5. Genes with multiple *de novo* **mutations in 125 congenital hydrocephalus case-parent trios (observed vs. expected)**. Observed and expected numbers of genes with greater than 1 *de novo* mutation in each variant category using a Monte Carlo simulation

	Observed	Expected	Enrichment	P-value
Total	5	0.948	5.27	**2.4 x 10^{-3}**
Synonymous	0	0.053	0	1
Missense	3	0.407	7.37	**7.4 x 10^{-3}**
Damaging (LOF + D-Mis)	3	0.074	40.81	**5.5 x 10^{-5}**
D-Missense	1	0.028	35.46	**0.028**
Loss of Function (LOF)	1	0.013	76.37	**0.013**

N: Number of case-parent trios, singleton cases, and familial forms (defined as greater than one patient in a nuclear family affected by congenital, primary hydrocephalus)

Supplementary Table 6. High brain expressed genes (Top Quartile) with multiple *de novo* **mutations in 125 congenital hydrocephalus case-parent trios (observed vs. expected).** Observed and expected numbers of high brain expressed genes (top 25% in the developing mouse brain) with greater than 1 *de novo* mutation in each variant category

	Observed	Expected	Enrichment	P-value
Total	4	0.26	15.5	**1.1 x 10^{-4}**
Synonymous	0	0.01	0	1
Missense	2	0.11	18.4	**5.4 x 10^{-3}**
Damaging (LOF + D-Mis)	3	0.02	151.7	**1.0 x 10^{-6}**
D-Missense	1	0.008	132.6	**7.5 x 10^{-3}**
Loss of Function (LOF)	1	0.004	279.5	**3.6 x 10^{-3}**

N: Number of case-parent trios, singleton cases, and familial forms (defined as greater than one patient in a nuclear family affected by congenital, primary hydrocephalus)

Supplementary Table 7. *De novo* mutations in TRIM71/LIN-41

Patient ID	Gene	Mutation	AA Change	Chr	Pos	Ref	Alt	gnomAD	ExAC	pLI	Polyphen	SIFT	MetaSVM	CADD	MPC (deleterious if >2)	Substitutions among 46 species
Hydro100-1	TRIM71 / LIN41	*De Novo* Missense	p.Arg796His	3	32933083	G	A	novel	novel	0.99	D	D	T	26.3	2.3	0
Hydro101-1	TRIM71 / LIN41	*De Novo* Missense	p.Arg608His	3	32932519	G	A	novel	novel	0.99	D	D	T	32	2.3	0
Hydro102-1	TRIM71 / LIN41	*De Novo* Missense	p.Arg608His	3	32932519	G	A	novel	novel	0.99	D	D	T	32	2.3	0
Hydro102-5	TRIM71 / LIN41	Transmitted Missense	p.Arg608His	3	32932519	G	A	novel	novel	0.99	D	D	T	32	2.3	0

Supplementary Table 8. *De novo* and damaging transmitted mutations in SMARCC1/BAF155

Patient ID	Gene	Mutation	Inheritance	AA Change	Chr	Pos	Ref	Alt	gnomAD	ExAC	pLI	Polyphen	SIFT	MetaSVM	CADD	MPC (deleterious if >2)	Substitutions among 46 species
Hydro105-1	SMARCC1 / BAF155	*D-Mis Mutation*	*De Novo*	p.His526Pro	3	47718267	T	G	novel	novel	1.00	D	D	D	28.9	2.6	0
Hydro106-3	SMARCC1 / BAF155	*Frameshift Deletion*	*De Novo*[1]	p.Lys891fs*4	3	47663806	T	-	novel	novel	1.00	.	.	.	.	.	0
Hydro107-1	SMARCC1 / BAF155	*Frameshift Insertion*	*Transmitted*	p.Leu534fs*28	3	47718254	-	GAGTCCCCACT	novel	novel	1.00	.	.	.	.	.	3
Hydro108-1	SMARCC1 / BAF155	*Stop Gain Mutation*	*Transmitted*	p.Lys179*	3	47777565	T	A	novel	novel	1.00	.	.	.	37	.	0
Hydro109-1	SMARCC1 / BAF155	*Frameshift Insertion*	*Unknown*[2]	p.Thr415fs*29	3	47730896	-	TT	novel	novel	1.00	.	.	.	.	.	0

[1] This *de novo* mutation was transmitted from an unaffected father (Hydro106-3) to an affected proband (Hydro106-1).

[2] This proband's biological father is deceased and biological mother is estranged.

Supplementary Table 9. *De novo* and damaging transmitted mutations in PTCH1

Patient ID	Gene	Mutation	Inheritance	AA Change	Chr	Pos	Ref	Alt	gnomAD	ExAC	pLI	Polyphen	SIFT	MetaSVM	CADD	MPC (deleterious if >2)	Substitutions among 46 species
Hydro103-1	PTCH1	*Frameshift Deletion*	*De Novo*	p.Met152fs*1	9	98248093	AAC	A	novel	novel	1.00	.	.	.	.	.	.
Hydro104-1	PTCH1	*Splice Donor*	*De Novo*	c.1503+3T>C	9	98239826	A	G	novel	novel	1.00	.	.	.	.	.	0
HydroFam100-1	PTCH1	*Frameshift Deletion*	*Autosomal Dominant*	p.Leu664fs*12	9	98231286	GTGCGGA	-	novel	novel	1.00	.	.	.	.	.	.

Supplementary Table 10. *De novo* mutations and transmitted mutations in PLOD2

Patient ID	Gene	Mutation	AA Change	Chr	Pos	Ref	Alt	gnomAD	ExAC	Missense Z-score	Polyphen	SIFT	MetaSVM	CADD	MPC (deleterious if >2)	Substitutions among 46 species
Hydro229-1	PLOD2	D-Mis	p.Arg473Gln	3	145796985	C	T	3.26E-05	3.78E-05	-1.00	D	T	D	25	0.20	1
Hydro245-1	PLOD2	D-Mis	p.Thr643Met	3	145789068	G	A	2.85E-05	2.84E-05	-1.00	D	D	D	28	0.42	1

Supplementary Table 12. Top GO Terms from GO ontology enrichment analysis for *PTCH1, TRIM71, SMARCC1*

GO Term	Description	P-value	FDR q-value	Enrichment	Genes
GO:0030850	prostate gland development	1.04×10^{-6}	1.57×10^{-2}	1,078.3	*SMARCC1, PTCH1*
GO:0001843	neural tube closure	5.11×10^{-5}	3.84×10^{-1}	160.3	*PTCH1, TRIM71*
GO:0060606	tube closure	5.39×10^{-5}	2.70×10^{-1}	156.1	*PTCH1, TRIM71*
GO:0035148	tube formation	1.28×10^{-4}	4.82×10^{-1}	101.4	*PTCH1, TRIM71*
GO:0060429	epithelium development	2.11×10^{-4}	6.34×10^{-1}	79.1	*PTCH1, TRIM71*
GO:0010157	response to chlorate	3.37×10^{-4}	8.46×10^{-1}	2,965.3	*PTCH1*
GO:0021997	neural plate axis specification	3.37×10^{-4}	7.25×10^{-1}	2,965.3	*PTCH1*
GO:0009957	epidermal cell fate specification	5.06×10^{-4}	9.52×10^{-1}	1,976.9	*PTCH1*
GO:0048732	gland development	5.39×10^{-4}	9.01×10^{-1}	49.4	*SMARCC1, PTCH1*
GO:0051254	positive regulation of RNA metabolic process	5.44×10^{-4}	8.18×10^{-1}	12.3	*SMARCC1, PTCH1, TRIM71*
GO:0048608	reproductive structure development	6.86×10^{-4}	9.40×10^{-1}	43.8	*SMARCC1, PTCH1*
GO:0060603	mammary gland duct morphogenesis	8.43×10^{-4}	1.00	1,186.1	*PTCH1*
GO:0045935	positive regulation of nucleobase-containing compound metabolic process	8.79×10^{-4}	1.00	10.4	*SMARCC1, PTCH1, TRIM71*

Supplementary Table 13. *De novo* enrichment analysis for neural tube defect genes in cases and controls

	Cases, N=125						Controls, N=1,789					
	Observed		Expected		Enrichment	*p*		Observed		Expected		Enrichment	*p*
	N	Rate	N	Rate				N	Rate	N	Rate		
Neural tube defect genes (N = 184)							**Neural tube defect genes (N = 184)**						
All	10	0.080	2	0.016	4.96	5.0×10^{-5}	All	26	0.015	27.8	0.016	0.94	0.66
Synonymous	1	0.008	0.6	0.005	1.69	0.45	Synonymous	7	0.004	8.1	0.005	0.87	0.69
Missense	6	0.048	1.3	0.010	4.74	2.0×10^{-3}	Missense	18	0.010	17.5	0.010	1.03	0.48
D-Mis	1	0.008	0.3	0.002	3.27	0.26	D-Mis	6	0.003	4.3	0.002	1.39	0.27
LOF	3	0.024	0.2	0.002	18.7	6.1×10^{-4}	LOF	1	0.001	2.3	0.001	0.44	0.90
Protein-Altering	9	0.072	1.4	0.011	6.32	1.9×10^{-5}	Protein-Altering	19	0.011	19.8	0.011	0.96	0.60
Damaging	4	0.032	0.5	0.004	8.58	1.4×10^{-3}	Damaging	7	0.004	6.6	0.004	1.06	0.49

Supplementary Table 14. *De novo* copy number variants in 125 case-parent trios with congenital hydrocephalus (CH)

Patient ID	Chr	Start	End	Length	Alteration	SQ score	# Exons	Genes in the region	1000G frequency
Hydro154-1	5	170814836	171100000	285164	Duplication	96	15	FGF18, NPM1	Novel
Hydro154-1	5	172068277	172342131	273854	Duplication	96	15	DUSP1, ERGIC1, LOC101928093, NEURL1B	0.002
Hydro134-1	6	170591818	170713790	121972	Deletion	95	18	DLL1, FAM120B	Novel
Hydro269-1	7	154735732	155604899	869167	Duplication	95	54	CNPY1, EN2, HTR5A, INSIG1, PAXIP1, RBM33, SHH	Novel
Hydro131-1	7	155599390	155755985	156595	Duplication	76	3	LOC389602, SHH	Novel
Hydro198-1	11	61731700	61734947	3247	Duplication	95	4	FTH1	Novel

References

Adle-Biassette, H., Saugier-Veber, P., Fallet-Bianco, C., Delezoide, A.-L., Razavi, F., Drouot, N., Bazin, A., Beaufrère, A.-M., Bessières, B., Blesson, S., *et al*. (2013). Neuropathological review of 138 cases genetically tested for X-linked hydrocephalus: evidence for closely related clinical entities of unknown molecular bases. Acta Neuropathologica 126, 427-442.

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nature methods 7, 248-249.

Aeschimann, F., Kumari, P., Bartake, H., Gaidatzis, D., Xu, L., Ciosk, R., and Großhans, H. LIN41 Post-transcriptionally Silences mRNAs by Two Distinct and Position-Dependent Mechanisms. Molecular Cell 65, 476-489.e474.

Al-Dosari, M.S., Al-Owain, M., Tulbah, M., Kurdi, W., Adly, N., Al-Hemidan, A., Masoodi, T.A., Albash, B., and Alkuraya, F.S. (2012). Mutation in MPDZ causes severe congenital hydrocephalus. Journal of Medical Genetics 50, 54.

Allache, R., Lachance, S., Guyot, M.C., De Marco, P., Merello, E., Justice, M.J., Capra, V., and Kibar, Z. (2014). Novel mutations in Lrp6 orthologs in mouse and human neural tube defects affect a highly dosage-sensitive Wnt non-canonical planar cell polarity pathway. Human molecular genetics 23, 1687-1699.

Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., *et al*. (2013). De novo mutations in epileptic encephalopathies. Nature 501, 217-221.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of molecular biology 215, 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research 25, 3389-3402.

Amlashi, S.F., Riffaud, L., Brassier, G., and Morandi, X. (2003). Nevoid basal cell carcinoma syndrome: relation with desmoplastic medulloblastoma in infancy. A population-based study and review of the literature. Cancer 98, 618-624.

Awadalla, P., Gauthier, J., Myers, R.A., Casals, F., Hamdan, F.F., Griffing, A.R., Cote, M., Henrion, E., Spiegelman, D., Tarabeux, J., *et al*. (2010). Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. Am J Hum Genet 87, 316-324.

Barak, T., Kwan, K.Y., Louvi, A., Demirbilek, V., Saygi, S., Tuysuz, B., Choi, M., Boyaci, H., Doerschner, K., Zhu, Y., *et al*. (2011). Recessive LAMC3 mutations cause malformations of occipital cortical development. Nature genetics 43, 590-594.

Bilguvar, K., Ozturk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoglu, D., Tuysuz, B., Caglayan, A.O., Gokben, S., *et al*. (2010). Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. Nature 467, 207-210.

Boivin, M.J., Kakooza, A.M., Warf, B.C., Davidson, L.L., and Grigorenko, E.L. (2015). Reducing neurodevelopmental disorders and disability through research and interventions. Nature 527, S155-160.

Boratyn, G.M., Schaffer, A.A., Agarwala, R., Altschul, S.F., Lipman, D.J., and Madden, T.L. (2012). Domain enhanced lookup time accelerated BLAST. Biology direct 7, 12.

Bret, P., and Chazal, J. (1995). Chronic ("normal pressure") hydrocephalus in childhood and adolescence. A review of 16 cases and reappraisal of the syndrome. Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 11, 687-691.

Briscoe, J., and Therond, P.P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol 14, 416-429.

Browning, S.R., and Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. American journal of human genetics 81, 1084-1097.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC bioinformatics 10, 421.

Cao, M., and Wu, J.I. (2015). Camk2a-Cre-Mediated Conditional Deletion of Chromatin Remodeler Brg1 Causes Perinatal Hydrocephalus. Neuroscience letters 597, 71-76.

Cappi, C., Brentani, H., Lima, L., Sanders, S.J., Zai, G., Diniz, B.J., Reis, V.N., Hounie, A.G., Conceicao do Rosario, M., Mariani, D., *et al*. (2016). Whole-exome sequencing in obsessive-compulsive disorder identifies rare mutations in immunological and neurodevelopmental pathways. Translational psychiatry 6, e764.

Cappi, C., Oliphant, M.E., Peter, Z., Zai, G., Sullivan, C.A., Gupta, A.R., Hoffman, E.J., Virdee, M., Willsey, A.J., Shavitt, R.G., *et al*. (2017). De novo damaging coding mutations are strongly associated with obsessive-compulsive disorder and overlap with autism. bioRxiv.

Carter, C.S., Vogel, T.W., Zhang, Q., Seo, S., Swiderski, R.E., Moninger, T.O., Cassell, M.D., Thedens, D.R., Keppler-Noreuil, K.M., Nopoulos, P., *et al*. (2012). Abnormal development of NG2+PDGFR-alpha+ neural progenitor cells leads to neonatal hydrocephalus in a ciliopathy mouse model. Nat Med 18, 1797-1804.

Carter, H., Douville, C., Stenson, P.D., Cooper, D.N., and Karchin, R. (2013). Identifying Mendelian disease genes with the variant effect scoring tool. BMC genomics 14 Suppl 3, S3.

Celen, C., Chuang, J.C., Luo, X., Nijem, N., Walker, A.K., Chen, F., Zhang, S., Chung, A.S., Nguyen, L.H., Nassour, I., *et al*. (2017). Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. eLife 6.

Chang, H.M., Martinez, N.J., Thornton, J.E., Hagan, J.P., Nguyen, K.D., and Gregory, R.I. (2012). Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. Nature communications 3, 923.

Chen, C.P., Lin, S.P., Wang, T.H., Chen, Y.J., Chen, M., and Wang, W. (2006). Perinatal findings and molecular cytogenetic analyses of de novo interstitial deletion of 9q (9q22.3-->q31.3) associated with Gorlin syndrome. Prenatal diagnosis 26, 725-729.

Chen, J., Lai, F., and Niswander, L. (2012). The ubiquitin ligase mLin41 temporally promotes neural progenitor cell maintenance through FGF signaling. Genes & development 26, 803-815.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the functional effect of amino acid substitutions and indels. PLoS One 7, e46688.

Chun, S., and Fay, J.C. (2009). Identification of deleterious mutations within three human genomes. Genome research 19, 1553-1561.

Cromer, M.K., Starker, L.F., Choi, M., Udelsman, R., Nelson-Williams, C., Lifton, R.P., and Carling, T. (2012). Identification of Somatic Mutations in Parathyroid Tumors Using Whole-Exome Sequencing. The Journal of Clinical Endocrinology & Metabolism 97, E1774-E1781.

Cuevas, E., Rybak-Wolf, A., Rohde, A.M., Nguyen, D.T.T., and Wulczyn, F.G. (2015). Lin41/Trim71 is essential for mouse development and specifically expressed in postnatal ependymal cells of the brain. Frontiers in Cell and Developmental Biology 3, 20.

Da, G., Lenkart, J., Zhao, K., Shiekhattar, R., Cairns, B.R., and Marmorstein, R. (2006). Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. Proceedings of the National Academy of Sciences of the United States of America 103, 2057-2062.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS computational biology 6, e1001025.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., *et al*. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43, 491-498.

Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., and Beroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic acids research 37, e67.

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., and Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Human molecular genetics 24, 2125-2137.

Duncan, S.A., Manova, K., Chen, W.S., Hoodless, P., Weinstein, D.C., Bachvarova, R.F., and Darnell, J.E., Jr. (1994). Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. Proceedings of the National Academy of Sciences of the United States of America 91, 7598-7602.

Ecsedi, M., and Grosshans, H. (2013). LIN-41/TRIM71: emancipation of a miRNA target. Genes & development 27, 581-589.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48.

Eggenschwiler, J.T., and Anderson, K.V. (2007). Cilia and developmental signaling. Annual review of cell and developmental biology 23, 345-373.

Eggenschwiler, J.T., Espinoza, E., and Anderson, K.V. (2001). Rab23 is an essential negative regulator of the mouse Sonic hedgehog signalling pathway. Nature 412, 194.

Ekici, A., Hilfinger, D., Jatzwauk, M., Thiel, C., Wenzel, D., Lorenz, I., Boltshauser, E., Goecke, T., Staatz, G., Morris-Rosendahl, D., *et al*. (2010). Disturbed Wnt Signalling due to a Mutation in CCDC88C Causes an Autosomal Recessive Non-Syndromic Hydrocephalus with Medial Diverticulum. Molecular Syndromology 1, 99-112.

Ellis, T., Smyth, I., Riley, E., Graham, S., Elliot, K., Narang, M., Kay, G.F., Wicking, C., and Wainwright, B. (2003). Patched 1 conditional null allele in mice. genesis 36, 158-161.

Ever, L., and Gaiano, N. (2005). Radial 'glial' progenitors: neurogenesis and signaling. Current opinion in neurobiology 15, 29-33.

Feng, W., Choi, I., Clouthier, D.E., Niswander, L., and Williams, T. (2013). The Ptch1(DL) mouse: a new model to study lambdoid craniosynostosis and basal cell nevus syndrome-associated skeletal defects. Genesis 51, 677-689.

Fischbach, G.D., and Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 68, 192-195.

Fromer, M., Moran, Jennifer L., Chambert, K., Banks, E., Bergen, Sarah E., Ruderfer, Douglas M., Handsaker, Robert E., McCarroll, Steven A., O'Donovan, Michael C., Owen, Michael J., *et al*. (2012). Discovery and Statistical Genotyping of Copy-Number Variation from Whole-Exome Sequencing Depth. American Journal of Human Genetics 91, 597-607.

Fromer, M., and Purcell, S.M. (2001). Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data. In Current protocols in human genetics (John Wiley & Sons, Inc.).

Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Rieder, M.J., Altshuler, D., Shendure, J., *et al*. (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493, 216-220.

Garber, M., Guttman, M., Clamp, M., Zody, M.C., Friedman, N., and Xie, X. (2009). Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics (Oxford, England) 25, i54-62.

Gavino, C., and Richard, S. (2011). Patched1 haploinsufficiency impairs ependymal cilia function of the quaking viable mice, leading to fatal hydrocephalus. Molecular and Cellular Neuroscience 47, 100-107.

Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., *et al*. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature 511, 344-347.

Goodrich, L.V., Johnson, R.L., Milenkovic, L., McMahon, J.A., and Scott, M.P. (1996). Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. Genes & development 10, 301-312.

Goodrich, L.V., Milenkovic, L., Higgins, K.M., and Scott, M.P. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants. Science (New York, NY) 277, 1109-1113.

Gorlin, R.J. (2004). Nevoid basal cell carcinoma (Gorlin) syndrome. Genetics in medicine : official journal of the American College of Medical Genetics 6, 530-539.

Gorlin , R.J., and Goltz , R.W. (1960). Multiple Nevoid Basal-Cell Epithelioma, Jaw Cysts and Bifid Rib. New England Journal of Medicine 262, 908-912.

Govaert, P., Oostra, A., Matthys, D., Vanhaesebrouck, P., and Leroy, J. (1991). How idiopathic is idiopathic external hydrocephalus? Developmental medicine and child neurology 33, 274-276.

Hahn, H., Wicking, C., Zaphiropoulous, P.G., Gailani, M.R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Unden, A.B., Gillies, S., *et al*. (1996). Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome. Cell 85, 841-851.

Harmacek, L., Watkins-Chow, D.E., Chen, J., Jones, K.L., Pavan, W.J., Salbaum, J.M., and Niswander, L. (2014). A Unique Missense Allele of BAF155, a Core BAF Chromatin Remodeling Complex Protein, Causes Neural Tube Closure Defects in Mice. Developmental neurobiology 74, 483-497.

Haverkamp, F., Wolfle, J., Aretz, M., Kramer, A., Hohmann, B., Fahnenstich, H., and Zerres, K. (1999). Congenital hydrocephalus internus and aqueduct stenosis: aetiology and implications for genetic counselling. European journal of pediatrics 158, 474-478.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proceedings of the National Academy of Sciences of the United States of America 106, 5181-5186.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., *et al*. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science (New York, NY) 350, 1262-1266.

Ioannidis, Nilah M., Rothstein, Joseph H., Pejaver, V., Middha, S., McDonnell, Shannon K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., *et al*. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. American Journal of Human Genetics 99, 877-885.

Ionita-Laza, I., McCallum, K., Xu, B.I.N., and Buxbaum, J. (2016). A Spectral Approach Integrating Functional Genomic Annotations for Coding and Noncoding Variants. Nature genetics 48, 214-220.

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.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., *et al*. (2012). De novo gene disruptions in children on the autistic spectrum. Neuron 74, 285-299.

Jagadeesh, K.A., Wenger, A.M., Berger, M.J., Guturu, H., Stenson, P.D., Cooper, D.N., Bernstein, J.A., and Bejerano, G. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nature genetics 48, 1581.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., *et al*. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nature genetics 49, 1593-1601.

Kahle, K.T., Kulkarni, A.V., Limbrick, D.D., Jr., and Warf, B.C. (2016). Hydrocephalus in children. The Lancet 387, 788-799.

Kamuro, K., and Tenokuchi, Y. (1993). Familial periventricular nodular heterotopia. Brain & development 15, 237-241.

Kanamoto, T. (2006). Cloning and regulation of the vertebrate homologue of lin-41 that functions as a heterochronic gene in Caenorhabditis elegans. Developmental dynamics 235, 1142-1149.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. Genome research 12, 656-664.

Kim, J.K., Huh, S.O., Choi, H., Lee, K.S., Shin, D., Lee, C., Nam, J.S., Kim, H., Chung, H., Lee, H.W., *et al*. (2001). Srg3, a mouse homolog of yeast SWI3, is essential for early embryogenesis and involved in brain development. Molecular and cellular biology 21, 7787-7795.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics 46, 310-315.

Kousi, M., and Katsanis, N. (2016). The Genetic Basis of Hydrocephalus. Annual review of neuroscience 39, 409-435.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.-X., *et al*. (2015). Excess of rare, inherited truncating mutations in autism. Nature genetics 47, 582-588.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols 4, 1073-1081.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., *et al*. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285-291.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England) 25, 1754-1760.

Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., *et al*. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. Nature 478, 476-482.

Loedige, I., Gaidatzis, D., Sack, R., Meister, G., and Filipowicz, W. (2013). The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. Nucleic acids research 41, 518-532.

Loedige, I., Jakob, L., Treiber, T., Ray, D., Stotz, M., Treiber, N., Hennig, J., Cook, K.B., Morris, Q., Hughes, T.R., *et al*. (2015). The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. Cell Rep 13, 1206-1220.

Loedige, I., Stotz, M., Qamar, S., Kramer, K., Hennig, J., Schubert, T., Löffler, P., Längst, G., Merkl, R., Urlaub, H., and Meister, G. (2014). The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. Genes & development 28, 749-764.

Madden, T.L., Tatusov, R.L., and Zhang, J. (1996). Applications of network BLAST server. Methods Enzymol 266, 131-141.

Maller Schulman, B.R., Liang, X., Stahlhut, C., DelConte, C., Stefani, G., and Slack, F.J. (2008). The let-7 microRNA target gene, Mlin41/Trim71 is required for mouse embryonic survival and neural tube closure. Cell cycle (Georgetown, Tex) 7, 3935-3942.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics (Oxford, England) 26, 2867-2873.

Mazzola, C.A., Choudhri, A.F., Auguste, K.I., Limbrick, D.D., Jr., Rogido, M., Mitchell, L., and Flannery, A.M. (2014). Pediatric hydrocephalus: systematic literature review and evidence-based guidelines. Part 2: Management of posthemorrhagic hydrocephalus in premature infants. Journal of neurosurgery Pediatrics 14 Suppl 1, 8-23.

McAllister, J.P., Guerra, M.M., Ruiz, L.C., Jimenez, A.J., Dominguez-Pinos, D., Sival, D., den Dunnen, W., Morales, D.M., Schmidt, R.E., Rodriguez, E.M., and Limbrick, D.D. (2017). Ventricular Zone Disruption in Human Neonates With Intraventricular Hemorrhage. Journal of neuropathology and experimental neurology 76, 358-375.

McAllister, J.P., Williams, M.A., Walker, M.L., Kestle, J.R.W., Relkin, N.R., Anderson, A.M., Gross, P.H., and Browd, S.R. (2015). An update on research priorities in hydrocephalus: overview of the third National Institutes of Health-sponsored symposium "Opportunities for Hydrocephalus Research: Pathways to Better Outcomes". Journal of neurosurgery 123, 1427-1438.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297-1303.

McPherson, E., and Clemens, M. (1997). Bruck syndrome (osteogenesis imperfecta with congenital joint contractures): review and report on the first North American case. American journal of medical genetics 70, 28-31.

Merkle, F.T., Tramontin, A.D., Garcia-Verdugo, J.M., and Alvarez-Buylla, A. (2004). Radial glia give rise to adult neural stem cells in the subventricular zone. Proceedings of the National Academy of Sciences of the United States of America 101, 17528-17532.

Milenkovic, L., Goodrich, L.V., Higgins, K.M., and Scott, M.P. (1999). Mouse patched1 controls body size determination and limb patterning. Development (Cambridge, England) 126, 4431-4440.

Mishra-Gorur, K., Caglayan, A.O., Schaffer, A.E., Chabu, C., Henegariu, O., Vonhoff, F., Akgumus, G.T., Nishimura, S., Han, W., Tu, S., *et al*. (2014). Mutations in KATNB1 cause complex cerebral malformations by disrupting asymmetrically dividing neural progenitors. Neuron 84, 1226-1239.

Mitschka, S., Ulas, T., Goller, T., Schneider, K., Egert, A., Mertens, J., Brustle, O., Schorle, H., Beyer, M., Klee, K., *et al*. (2015). Co-existence of intact stemness and priming of neural differentiation programs in mES cells lacking Trim71. Sci Rep 5, 11126.

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R., and Schaffer, A.A. (2008). Database indexing for production MegaBLAST searches. Bioinformatics (Oxford, England) 24, 1757-1764.

Munch, T.N., Rostgaard, K., Rasmussen, M.L., Wohlfahrt, J., Juhler, M., and Melbye, M. (2012). Familial aggregation of congenital hydrocephalus in a nationwide cohort. Brain : a journal of neurology 135, 2409-2415.

Narayanan, R., Pirouz, M., Kerimoglu, C., Pham, L., Wagener, R.J., Kiszka, K.A., Rosenbusch, J., Seong, R.H., Kessel, M., Fischer, A., *et al*. (2015). Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. Cell Rep 13, 1842-1854.

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.

Ng, P.C., and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome research 11, 863-874.

Ng, P.C., and Henikoff, S. (2002). Accounting for human polymorphisms predicted to affect protein function. Genome research 12, 436-446.

Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research 31, 3812-3814.

Ng, P.C., and Henikoff, S. (2006). Predicting the effects of amino acid substitutions on protein function. Annual review of genomics and human genetics 7, 61-80.

Nguyen, D.T.T., Richter, D., Michel, G., Mitschka, S., Kolanus, W., Cuevas, E., and Wulczyn, F.G. (2017). The ubiquitin ligase LIN41/TRIM71 targets p53 to antagonize cell death and differentiation pathways during stem cell differentiation. Cell death and differentiation 24, 1063-1078.

Nguyen, H., Sokpor, G., Pham, L., Rosenbusch, J., Stoykova, A., Staiger, J.F., and Tuoc, T. (2016). Epigenetic regulation by BAF (mSWI/SNF) chromatin remodeling complexes is indispensable for embryonic development. Cell cycle (Georgetown, Tex) 15, 1317-1324.

Noctor, S.C., Martinez-Cerdeno, V., Ivic, L., and Kriegstein, A.R. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nature neuroscience 7, 136-144.

O'Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C., *et al*. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nature genetics 43, 585-589.

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*. (2012). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science (New York, NY) 338, 1619-1622.

Oreskovic, D., and Klarica, M. (2011). Development of hydrocephalus and classical hypothesis of cerebrospinal fluid hydrodynamics: facts and illusions. Progress in neurobiology 94, 238-258.

Palma, V., Lim, D.A., Dahmane, N., Sánchez, P., Brionne, T.C., Herzberg, C.D., Gitton, Y., Carleton, A., Álvarez-Buylla, A., and Altaba, A.R.i. (2005). Sonic hedgehog controls stem cell behavior in the postnatal and adult brain. Development (Cambridge, England) 132, 335-344.

Palma, V., and Ruiz i Altaba, A. (2004). Hedgehog-GLI signaling regulates the behavior of cells with stem cell properties in the developing neocortex. Development (Cambridge, England) 131, 337-345.

Petersen, P.H., Zou, K., Hwang, J.K., Jan, Y.N., and Zhong, W. (2002). Progenitor cell maintenance requires numb and numblike during mouse neurogenesis. Nature 419, 929-934.

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome research 20, 110-121.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics 38, 904.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81, 559-575.

Quang, D., Chen, Y., and Xie, X. (2015). DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics (Oxford, England) 31, 761-763.

Reinhart, B.J. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature (London) 403, 901-906.

Rekate, H.L. (2008). The definition and classification of hydrocephalus: a personal recommendation to stimulate debate. Cerebrospinal Fluid Research 5, 2-2.

Reva, B., Antipin, Y., and Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic acids research 39, e118.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature biotechnology 29, 24-26.

Robinson, S. (2012). Neonatal posthemorrhagic hydrocephalus from prematurity: pathophysiology and current treatment concepts. Journal of neurosurgery Pediatrics 9, 242-258.

Rosenthal, A., Jouet, M., and Kenwrick, S. (1992). Aberrant splicing of neural cell adhesion molecule L1 mRNA in a family with X-linked hydrocephalus. Nature genetics 2, 107-112.

Ruderfer, D.M., Hamamsy, T., Lek, M., Karczewski, K.J., Kavanagh, D., Samocha, K.E., Exome Aggregation, C., Daly, M.J., MacArthur, D.G., Fromer, M., and Purcell, S.M. (2016). Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. Nature genetics 48, 1107-1111.

Samocha, K.E., Kosmicki, J.A., Karczewski, K.J., O'Donnell-Luria, A.H., Pierce-Hoffman, E., MacArthur, D.G., Neale, B.M., and Daly, M.J. (2017). Regional missense constraint improves variant deleteriousness prediction. bioRxiv.

Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnstrom, K., Mallick, S., Kirby, A., *et al*. (2014). A framework for the interpretation of de novo mutation in human disease. Nature genetics 46, 944-950.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., *et al*. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485, 237-241.

Sato, O., Takei, F., and Yamada, S. (1994). Hydrocephalus: is impaired cerebrospinal fluid circulation only one problem involved? Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 10, 151-155.

Satterstrom, F.K., Walters, R.K., Singh, T., Wigdor, E.M., Lescai, F., Demontis, D., Kosmicki, J.A., Grove, J., Stevens, C., Bybjerg-Grauholm, J., *et al*. (2018). ASD and ADHD have a similar burden of rare protein-truncating variants. bioRxiv.

Schaeren-Wiemers, N., and Gerfin-Moser, A. (1993). A single protocol to detect transcripts of various types and expression levels in neural tissue and cultured cells: in situ hybridization using digoxigenin-labelled cRNA probes. Histochemistry 100, 431-440.

Schulman, B.R., Esquela-Kerscher, A., and Slack, F.J. (2005). Reciprocal expression of lin-41 and the microRNAs let-7 and mir-125 during mouse embryogenesis. Developmental dynamics : an official publication of the American Association of Anatomists 234, 1046-1054.

Schwarz, J.M., Rodelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. Nature methods 7, 575-576.

Shaheen, R., Sebai, M.A., Patel, N., Ewida, N., Kurdi, W., Altweijri, I., Sogaty, S., Almardawi, E., Seidahmed, M.Z., Alnemri, A., *et al*. (2017). The genetic landscape of familial congenital hydrocephalus. Annals of neurology 81, 890-897.

Shannon, C.N., Simon, T.D., Reed, G.T., Franklin, F.A., Kirby, R.S., Kilgore, M.L., and Wellons, J.C., 3rd (2011). The economic impact of ventriculoperitoneal shunt failure. Journal of neurosurgery Pediatrics 8, 593-599.

Sheen, V.L., Basel-Vanagaite, L., Goodman, J.R., Scheffer, I.E., Bodell, A., Ganesh, V.S., Ravenscroft, R., Hill, R.S., Cherry, T.J., Shugart, Y.Y., *et al*. (2004). Etiological heterogeneity of familial periventricular heterotopia and hydrocephalus. Brain & development 26, 326-334.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic acids research 29, 308-311.

Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L., Edwards, K.J., Day, I.N., and Gaunt, T.R. (2013). Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 34, 57-65.

Shihab, H.A., Gough, J., Mort, M., Cooper, D.N., Day, I.N., and Gaunt, T.R. (2014). Ranking non-synonymous single nucleotide polymorphisms based on disease concepts. Human genomics 8, 11.

Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N.M., Gaunt, T.R., and Campbell, C. (2015). An integrative approach to predicting the functional effects of non-coding and coding sequence variation. Bioinformatics (Oxford, England) 31, 1536-1543.

Shikata, Y., Okada, T., Hashimoto, M., Ellis, T., Matsumaru, D., Shiroishi, T., Ogawa, M., Wainwright, B., and Motoyama, J. (2011). Ptch1-mediated dosage-dependent action of Shh signaling regulates neural progenitor development at late gestational stages. Dev Biol 349, 147-159.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., *et al*. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research 15, 1034-1050.

Simon, T.D., Riva-Cambrin, J., Srivastava, R., Bratton, S.L., Dean, J.M., and Kestle, J.R. (2008). Hospital care for children with hydrocephalus in the United States: utilization, charges, comorbidities, and deaths. Journal of neurosurgery Pediatrics 1, 131-137.

Sippl, M.J. (1993). Recognition of errors in three-dimensional structures of proteins. Proteins 17, 355-362.

Slack, F.J. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Molecular cell 5, 659-669.

Slack, F.J., and Ruvkun, G. (1998). A novel repeat domain that is often associated with RING finger and B-box motifs. Trends in biochemical sciences 23, 474-475.

Slavotinek, A., Kaylor, J., Pierce, H., Cahr, M., DeWard, Stephanie J., Schneidman-Duhovny, D., Alsadah, A., Salem, F., Schmajuk, G., and Mehta, L. (2015). CRB2 Mutations Produce a Phenotype Resembling Congenital Nephrosis, Finnish Type, with Cerebral Ventriculomegaly and Raised Alpha-Fetoprotein. The American Journal of Human Genetics 96, 162-169.

States, D.J., and Gish, W. (1994). Combined use of sequence similarity and codon bias for coding region identification. Journal of computational biology : a journal of computational molecular cell biology 1, 39-50.

Stuart, B.D., Choi, J., Zaidi, S., Xing, C., Holohan, B., Chen, R., Choi, M., Dharwadkar, P., Torres, F., Girod, C.E., *et al*. (2015). Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. Nature genetics 47, 512-517.

Study, D.D.D. (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature 519, 223-228.

Study, D.D.D. (2017). Prevalence and architecture of de novo mutations in developmental disorders. Nature 542, 433-438.

Sudmant, P.H., Rausch, T., Gardner, E.J., Handsaker, R.E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Fritz, M.H.-Y., *et al*. (2015). An integrated map of structural variation in 2,504 human genomes. Nature 526, 75-81.

Svärd, J., Rozell, B., Toftgård, R., and Teglund, S. (2009). Tumor suppressor gene co-operativity in compound Patched1 and Suppressor of fused heterozygous mutant mice. Molecular carcinogenesis 48, 408-419.

Takahashi, M., and Osumi, N. (2002). Pax6 regulates specification of ventral neurone subtypes in the hindbrain by establishing progenitor domains. Development (Cambridge, England) 129, 1327-1338.

The Genomes Project, C. (2015). A global reference for human genetic variation. Nature 526, 68-74.

Timberlake, A.T., Choi, J., Zaidi, S., Lu, Q., Nelson-Williams, C., Brooks, E.D., Bilguvar, K., Tikhonova, I., Mane, S., Yang, J.F., *et al*. (2016). Two locus inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. eLife 5.

Timberlake, A.T., Furey, C.G., Choi, J., Nelson-Williams, C., Analysis, Y.C.f.G., Loring, E., Galm, A., Kahle, K.T., Steinbacher, D.M., Larysz, D., *et al*. (2017). De novo mutations in inhibitors of Wnt, BMP, and Ras/ERK signaling pathways in non-syndromic midline craniosynostosis. Proceedings of the National Academy of Sciences.

Tissir, F., Qu, Y., Montcouquiol, M., Zhou, L., Komatsu, K., Shi, D., Fujimori, T., Labeau, J., Tyteca, D., Courtoy, P., *et al*. (2010). Lack of cadherins Celsr2 and Celsr3 impairs ependymal ciliogenesis, leading to fatal hydrocephalus. Nature neuroscience 13, 700-707.

Tully, H.M., and Dobyns, W.B. (2014). Infantile hydrocephalus: a review of epidemiology, classification and causes. European journal of medical genetics 57, 359-368.

Tully, H.M., Ishak, G.E., Rue, T.C., Dempsey, J.C., Browd, S.R., Millen, K.J., Doherty, D., and Dobyns, W.B. (2016). Two Hundred Thirty-Six Children With Developmental Hydrocephalus: Causes and Clinical Consequences. Journal of child neurology 31, 309-320.

Tuoc, T.C., Boretius, S., Sansom, S.N., Pitulescu, M.E., Frahm, J., Livesey, F.J., and Stoykova, A. (2013a). Chromatin regulation by BAF170 controls cerebral cortical size and thickness. Developmental cell 25, 256-269.

Tuoc, T.C., Narayanan, R., and Stoykova, A. (2013b). BAF chromatin remodeling complex: cortical size regulation and beyond. Cell cycle (Georgetown, Tex) 12, 2953-2959.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., *et al*. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Current protocols in bioinformatics 43, 11.10.11-33.

Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes & development 18, 132-137.

Wang, C., Pan, Y., and Wang, B. (2010a). Suppressor of fused and Spop regulate the stability, processing and function of Gli2 and Gli3 full-length activators but not their repressors. Development (Cambridge, England) 137, 2001-2009.

Wang, K., Li, M., and Hakonarson, H. (2010b). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 38, e164.

Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B.R., and Crabtree, G.R. (1996). Diversity and specialization of mammalian SWI/SNF complexes. Genes & development 10, 2117-2130.

Ware, J.S., Samocha, K.E., Homsy, J., and Daly, M.J. (2015). Interpreting de novo Variation in Human Disease Using denovolyzeR. Current protocols in human genetics 87, 7.25.21-15.

Wetmore, C., Eberhart, D.E., and Curran, T. (2000). The normal patched allele is expressed in medulloblastomas from mice with heterozygous germ-line mutation of patched. Cancer research 60, 2239-2246.

Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P., and Deutschbauer, A. (2015). Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. mBio 6.

Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research 35, W407-W410.

Willsey, A.J., Fernandez, T.V., Yu, D., King, R.A., Dietrich, A., Xing, J., Sanders, S.J., Mandell, J.D., Huang, A.Y., Richer, P., *et al*. (2017). De Novo Coding Variants Are Strongly Associated with Tourette Disorder. Neuron 94, 486-499.e489.

Worringer, K.A., Rand, T.A., Hayashi, Y., Sami, S., Takahashi, K., Tanabe, K., Narita, M., Srivastava, D., and Yamanaka, S. (2014). The let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. Cell stem cell 14, 40-52.

Xu, B., Roos, J.L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., Gogos, J.A., and Karayiorgou, M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nature genetics 43, 864-868.

Yan, Z., Wang, Z., Sharova, L., Sharov, A.A., Ling, C., Piao, Y., Aiba, K., Matoba, R., Wang, W., and Ko, M.S. (2008). BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. Stem cells (Dayton, Ohio) 26, 1155-1165.

Yao, B., Christian, K.M., He, C., Jin, P., Ming, G.L., and Song, H. (2016). Epigenetic mechanisms in neurogenesis. Nat Rev Neurosci 17, 537-549.

Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. Journal of computational biology : a journal of computational molecular cell biology 11, 377-394.

Yu, G., Yang, Y., and Tian, G. (2010). Expressing and characterization of mLIN-41 in mouse early embryos and adult muscle tissues. Journal of molecular histology 41, 295-305.

Yung, Y.C., Mutoh, T., Lin, M.-E., Noguchi, K., Rivera, R.R., Choi, J.W., Kingsbury, M.A., and Chun, J. (2011). Lysophosphatidic Acid Signaling May Initiate Fetal Hydrocephalus. Science Translational Medicine 3, 99ra87-99ra87.

Zega, K., Jovanovic, V.M., Vitic, Z., Niedzielska, M., Knaapi, L., Jukic, M.M., Partanen, J., Friedel, R.F., Lang, R., and Brodski, C. (2017). Dusp16 Deficiency Causes Congenital Obstructive Hydrocephalus and Brain Overgrowth by Expansion of the Neural Progenitor Pool. Frontiers in molecular neuroscience 10, 372.

Zhang, J., and Madden, T.L. (1997). PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. Genome research 7, 649-656.

Zhang, J., Williams, M.A., and Rigamonti, D. (2006). Genetics of human hydrocephalus. Journal of Neurology 253, 1255-1266.

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. Journal of computational biology : a journal of computational molecular cell biology 7, 203-214.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Whole exome sequencing data from CH trios (n = 125)	This paper	Accession no. phs000744 http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000744
Whole exome sequencing data from SSC control trios	Iossifov et al., 2014	NDAR: DOI:10.15154/1149697 https://ndar.nih.gov/study.html?id=352
Software and Algorithms		
Genome Analysis Tool Kit (GATK)	DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2013	https://software.broadinstitute.org/gatk/best-practices/
BWA-mem	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
Annovar	Wang et al., 2010	http://annovar.openbioinformatics.org/en/latest/
PLINK/SEQ	Fromer et al., 2014	https://atgu.mgh.harvard.edu/plinkseq/
Other		
1000 Genomes GRCh37 h19 genome build	1000 Genomes Project	http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/human_g1k_v37.fasta.gz
RefSeq hg19 gene annotation	UCSC Genome Browser	http://genome.ucsc.edu/cgi-bin/hgTables?command=start
Intervals file for IDT xGen v1.0	Integrated DNA Technologies	https://www.idtdna.com/pages/products/next-generation-sequencing/hybridization-capture/lockdown-panels/xgen-exome-research-panel