OXFORD

ORIGINAL ARTICLE

HCFC1 loss-of-function mutations disrupt neuronal and neural progenitor cells of the developing brain

Lachlan A. Jolly[1,2], Lam Son Nguyen[4], Deepti Domingo[3], Ying Sun[1,2], Simon Barry[1], Miroslava Hancarova[5], Pavlina Plevova[6], Marketa Vlckova[5], Marketa Havlovicova[5], Vera M. Kalscheuer[7], Claudio Graziano[8], Tommaso Pippucci[8], Elena Bonora[8], Zdenek Sedlacek[5] and Jozef Gecz[1,2,3,*]

[1]School of Paediatrics and Reproductive Health, [2]Robinson Research Institute and [3]School of Molecular and Biomedical Sciences, University of Adelaide, Adelaide 5000, Australia, [4]INSERM UMR 1163, Laboratory of Molecular and Pathophysiological Bases of Cognitive Disorders, Paris Descartes—Sorbonne Paris Cité University, Imagine Institute, Necker-Enfants Malades Hospital, 75015 Paris, France, [5]Department of Biology and Medical Genetics, Charles University 2nd Faculty of Medicine and University Hospital Motol, Prague, Czech Republic, [6]Department of Medical Genetics, University Hospital Ostrava, tr. 17. listopadu 1790, 708 52 Ostrava, Czech Republic, [7]Department of Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Ihnestrasse 73, D-14195 Berlin, Germany and [8]Unit of Medical Genetics, Department of Medical and Surgical Sciences, DIMEC, St.Orsola-Malpighi Hospital, University of Bologna, Bologna 40138, Italy

*To whom correspondence should be addressed. Tel: +61 881616339; Fax: +61 881617342; Email: jozef.gecz@adelaide.edu.au

Abstract

Both gain- and loss-of-function mutations have recently implicated *HCFC1* in neurodevelopmental disorders. Here, we extend our previous *HCFC1* over-expression studies by employing short hairpin RNA to reduce the expression of *Hcfc1* in embryonic neural cells. We show that in contrast to over-expression, loss of Hcfc1 favoured proliferation of neural progenitor cells at the expense of differentiation and promoted axonal growth of post-mitotic neurons. To further support the involvement of *HCFC1* in neurological disorders, we report two novel *HCFC1* missense variants found in individuals with intellectual disability (ID). One of these variants, together with three previously reported *HCFC1* missense variants of unknown pathogenicity, were functionally assessed using multiple cell-based assays. We show that three out of the four variants tested result in a partial loss of HCFC1 function. While over-expression of the wild-type *HCFC1* caused reduction in HEK293T cell proliferation and axonal growth of neurons, these effects were alleviated upon over-expression of three of the four *HCFC1* variants tested. One of these partial loss-of-function variants disrupted a nuclear localization sequence and the resulting protein displayed reduced ability to localize to the cell nucleus. The other two variants displayed negative effects on the expression of the HCFC1 target gene *MMACHC*, which is responsible for the metabolism of cobalamin, suggesting that these individuals may also be susceptible to cobalamin deficiency. Together, our work identifies plausible cellular consequences of missense *HCFC1* variants and identifies likely and relevant disease mechanisms that converge on embryonic stages of brain development.

Introduction

Intellectual disability (ID) affects ~2–3% of the population and is considered one of the most genetically heterogeneous human disorders (1). Contribution of genes on the X chromosome (i.e. X-linked ID; XLID) has been the best characterized with over 100 genes implicated thus far. We have recently reported a non-coding regulatory mutation in an X-linked gene, *HCFC1* (OMIM 3 00 019), as the likely cause of mild non-syndromic ID in the large X-linked family MRX3 (2). HCFC1 is a transcriptional co-regulator with many important functions in cell proliferation and mitochondrial biogenesis (3–6). Recent data suggest that HCFC1 containing transcriptional complexes may regulate more than a quarter of all human promoters (7). The non-coding regulatory mutation in the 5′ untranslated region (UTR) of *HCFC1* abolished a DNA-binding site for the transcriptional regulator YY1, which subsequently caused de-repression of *HCFC1* gene expression in patient cell lines (2). To model the consequences of this change, over-expression of *HCFC1* was shown to alter the behaviour of embryonic neural cells, namely promoting differentiation of neural progenitor cells (NPCs) and reducing the neurite growth of *ex vivo* cultured hippocampal neurons. Furthermore, consistent with these newly identified roles for *HCFC1* and those previously established, transcriptome analysis of patient cell lines highlighted deregulation of genes that were important for embryonic brain growth, transcriptional regulation, and mitochondrial function (2,8). Additionally, three unrelated families containing male individuals with ID and different *HCFC1* missense changes were identified (2). Whether these additional variants were pathogenic, or simply low-frequency benign single-nucleotide polymorphisms (SNPs), was not established. Subsequently, multiple *HCFC1* missense mutations, affecting the N-terminal Kelch domain of the HCFC1 protein, have been identified in a cohort of patients with X-linked Cobalamin type C (CblC)-like disorder, which was named CblX (9,10) (OMIM: 3 09 541). CblC (OMIM: 2 77 400) is a metabolic disorder most frequently caused by mutations in *MMACHC* (OMIM: 6 09 831), which encodes an enzyme required for metabolism of cobalamin (11). HCFC1 was shown to bind to the promoter of *MMACHC* and knockdown of *HCFC1* in cultured cells led to a loss of *MMACHC* expression. Similarly and most importantly, *MMACHC* expression was significantly reduced also in skin fibroblasts of two CblX patients tested (9). Thus, loss of HCFC1 function was demonstrated as the cause of CblX. Both CblC and CblX patients display neurological impairment; however, CblX-related neurological phenotype is much more severe and includes intractable epilepsy, brain malformations and severe cognitive impairment (9). Together, these findings suggest that in addition to over-expression of *HCFC1,* also loss-of-function *HCFC1* mutations cause disruptions to normal brain development and likely involve multiple mechanisms.

In this study, we extend our original investigations of *HCFC1* over-expression by investigating the effect that loss-of-function *HCFC1* mutations have on embryonic brain development. We also report two novel *HCFC1* variants identified in two unrelated patients with ID and collate available clinical findings. Finally, we use multiple cell-based assays to address the functional consequence of these and also previously reported *HCFC1* variants. Our findings shed further light into the cellular mechanisms as well as the clinical consequences of *HCFC1*-related neurological disorders.

Results

Reduction of Hcfc1 stimulates neuronal growth

To model the underlying neural cell pathology of *Hcfc1* loss-of-function mutations, and to compare it with established mechanisms behind gain-of-function mutations, we developed short hairpin RNA (shRNA) lentiviral reagents to knock down *Hcfc1* expression. We identified one shRNA trigger sequence (T2) that targeted the untranslated region (UTR) of *Hcfc1*, that modestly knockdown *Hcfc1* expression by about half when tested in NIH3T3 cells (Supplementary Material, Fig. S1). As *HCFC1* gain-of-function is known to reduce the axonal growth and neurite arborization of hippocampal neurons, we asked what would be the effect of loss of function in these assays. We isolated post-mitotic hippocampal neurons from embryonic day 18.5 (E18.5) mice embryos and plated them *in vitro*. Following 12 h growth, we transduced the cells with lentiviral particles encoding enhanced green fluorescent protein (EGFP) and either a control shRNA sequence (sequence against luciferase that does not target mammalian genes) or a *Hcfc1* shRNA (T2; Hcfc1[shRNA]). Following transduction, neurons were allowed to grow until 4 days *in vitro*, subsequently fixed and immunofluorescently labelled against the marker proteins MAP2 and TAU1 to identify dendrites and axons, respectively (Fig. 1A). Next we conducted morphometric analysis on *EGFP* expressing (i.e. transduced) neurons. Compared with neurons transduced with control lentiviral particles, neurons transduced with *Hcfc1*[shRNA] viral particles displayed a 33% increase in axonal length and similar increases in axonal and dendritic (and hence total) termini number, which reports on the degree of neuronal arborization (Fig. 1B and C). Re-expression of *HCFC1* could rescue this axonal defect that confirms the specificity of the *Hcfc1*[shRNA] (see below). Thus, reduction of *Hcfc1* expression also affected hippocampal neuronal growth and displayed the opposite effect to that previously described for *HCFC1* over-expression (2).

### Loss of Hcfc1 results in expansion of neural progenitor cells

Promotion of NPC differentiation was also observed when *HCFC1* was over-expressed to model the effect of the gain-of-function *HCFC1* mutation (2). Thus, we investigated whether these cells would also be sensitive to loss of *Hcfc1* function. We isolated NPCs from the E18.5 embryonic dorsal cortex and grew them as non-adherent neurospheres. We transduced the cells with control or *Hcfc1*[shRNA] lentiviral particles, and achieved transduction rates of ~70% (data not shown). In the first instance, we created purified transduced cultures using fluorescence-activated cell sorting (FACS) and subjected purified cells to a cell proliferation assay. Cells were plated at an equivalent initial density and allowed to grow for 6 days during which cell growth in the cultures was assayed at Days 0, 3 and 6 (Cell Titre 96AQueous Assay; Promega). When seeded at low (clonal) density, cell growth was linear in both control and *Hcfc1*[shRNA] cultures (linear regression of line of best fit: $R^2 = 0.984$ and 0.997, respectively); however, the rate of growth in *Hcfc1*[shRNA] cultures was higher (slope of linear fit; control = 0.56, *Hcfc1*[shRNA] = 1.12), resulting in a 36 and 65% increase in cell numbers at Days 3 and 6 of culture, respectively (Fig. 2A). Similar results were obtained when cells were cultured at higher initial densities; however, cell growth plateaued at later stages, perhaps due to nutrient growth restriction (Supplementary Material, Fig. S2). We reasoned this increase in cell numbers may result from (i) an inhibition of differentiation even under proliferative culture conditions (i.e. in the presence of the epidermal growth factor; EGF), (ii) reduction in apoptosis or (iii) increased progenitor proliferation. To address these alternatives, we first collected RNA from the purified cultures and applied qRT-PCR to discover the expression levels of several cell type-specific marker genes (Fig. 2B). We saw no statistically significant
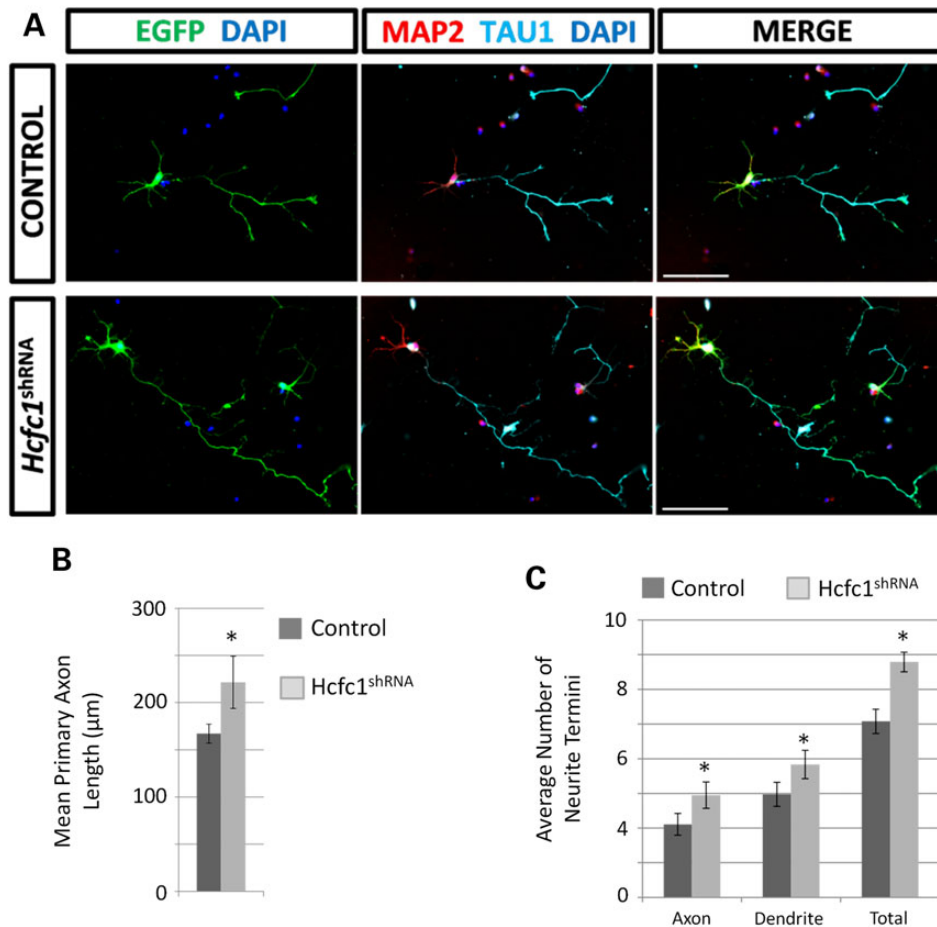
Figure 1. Loss of *Hcfc1* causes enhanced hippocampal neuronal growth *in vitro*. Isolated primary hippocampal neurons were cultured following transduction with lentiviral particles encoding either an inert shRNA sequence (against luciferase; control) or *Hcfc1* shRNA (Hcfc1^{shRNA}). (**A**) Representative immunofluorescent images of control and Hcfc1^{shRNA} neurons at Day 4 of *in vitro* growth. Bars = 100 μm. Transduced cells express *EGFP* (green). Cultures were co-stained using antibodies that label dendritic (MAP2; red) and axonal (TAU1; cyan) neuronal structures. (**B**) Quantification of primary axonal length. (**C**) Quantification of neurite termini. *$P < 0.05$ by Student's *t*-test.

difference in the expression of the early neuroepithelia marker *Sox1*, the cortical NPC marker *Pax6*, the radial glia cell marker *BLBP* nor the pan NPC marker *Nestin*. Likewise, we saw no significant differences in the terminally differentiated cell-type marker genes, namely the early neuronal marker *βIII-tubulin*, the astrocyte marker *GFAP* and the oligodendrocyte marker *CNPase*. While not reaching statistical significance, we did note a trend for elevated expression of marker genes for progenitor cell types and a reciprocal decrease in the expression of terminally differentiated cell-type marker genes in *Hcfc1*^{shRNA} cultures (Fig. 2B). To more definitively identify and quantify the numbers of progenitor cells and monitor their behaviours at single-cell resolution, we employed immunofluorescence. We dissociated neurosphere cells (from parallel non-purified cultures) and plated them onto a poly-l-lysine substrate in the presence of EGF. Only transduced cells (i.e. those expressing EGFP) were analysed. The percentage of transduced cells of NPC identity was reported by the co-expression of *EGFP* and *Pax6*. We identified a significant 20% increase in Pax6-positive NPCs in *Hcfc1*^{shRNA} cultures (Fig. 2C and Supplementary Material, Fig. S3A). We next asked whether apoptotic rates within the cultures were affected by reduction of *Hcfc1*; however, we were unable to identify any differences in the number of transduced cells antigenic to an activated caspase3 antibody, a hallmark of cells undergoing cell death (Fig. 2D and Supplementary Material, Fig. S3B). Next, we reasoned

that if more NPCs were present in the *Hcfc1*^{shRNA} cultures, then a higher percentage of transduced cells will be undergoing mitosis at any given time, so we used an antibody against phospho-histone3, a specific nuclear mark of cells at mitosis. Transduced cells in *Hcfc1*^{shRNA} cultures were found to have a 34% increase in cells undergoing mitosis compared with controls (Fig. 2E and Supplementary Material, Fig. S3C). Together the results suggest that under proliferative conditions, transduced NPCs in *Hcfc1*^{shRNA} cultures were more likely to remain in the proliferative progenitor state. We further explored the possibility that transduced cells in *Hcfc1*^{shRNA} cultures might be conducive to NPC maintenance under conditions that promote differentiation. We thus plated control and *Hcfc1*^{shRNA}-transduced cells onto poly-l-lysine substrate as described above but removed growth factor so as to promote differentiation. We allowed the cells to differentiate for 3 days before staining them immunofluorescently using antibodies against cell type-specific marker proteins Pax6, βIII-tubulin, GFAP and CNPase, and counted their abundance within the cultures (Fig. 2F and Supplementary Material, Fig. S4A–C). Under these conditions, transduced cells in the *Hcfc1*^{shRNA} cultures had 25% more Pax6-expressing NPCs, and 34% less differentiated cell types compared with transduced cells in control cultures. This reduction in differentiation was found to be significant across all three neural cell lineages. In aggregate, our data suggest that reduction of *Hcfc1* in NPCs promotes the cell cycling
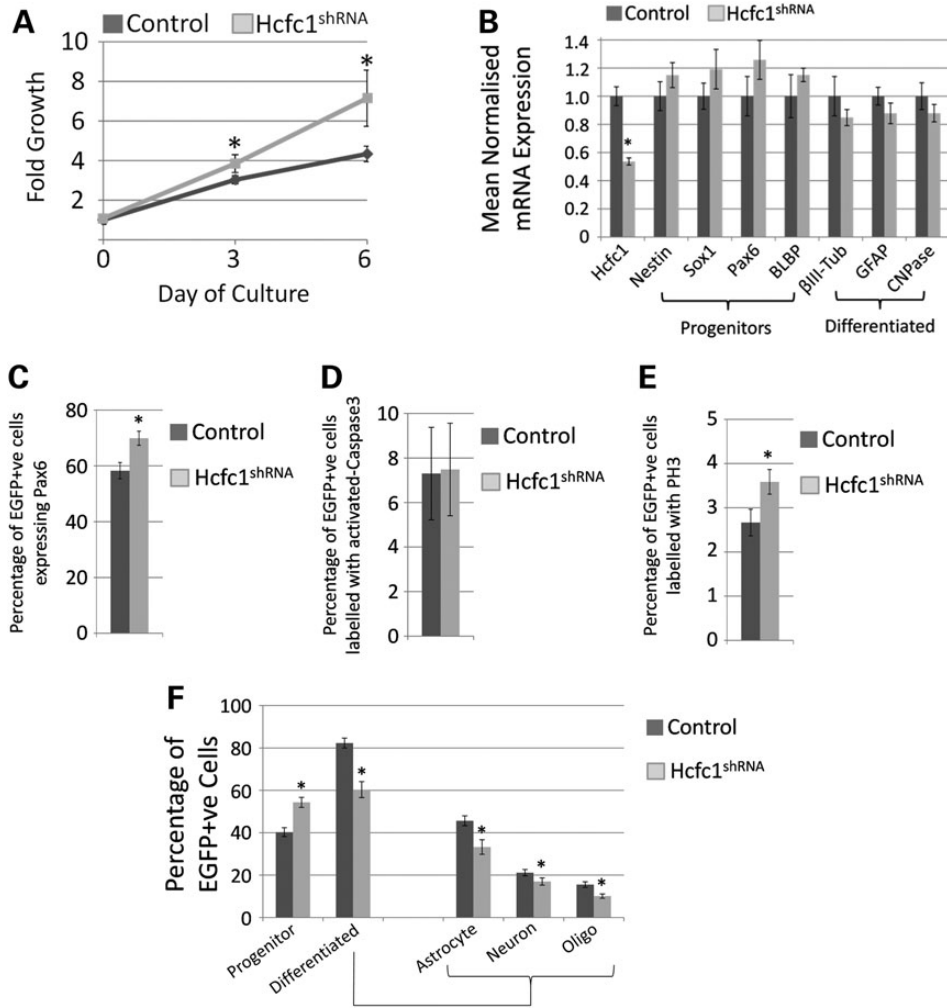
Figure 2. Loss of *Hcfc1* causes increased growth kinetics and reduced differentiation of NPCs. Isolated primary cortical NPCs grown *in vitro* as neurospheres were transduced with lentiviral particles encoding either inert shRNA sequences (Control) or *Hcfc1* targeted shRNA (*Hcfc1*^shRNA^). (**A** and **B**) Transduced cells were purified using FACS prior to assays. (A) Cell proliferation assay conducted in the presence of EGF at Days 0, 3 and 6 following passage. Cells were plated at an initial low clonal density of 1×10^4 cells. (B) qRT–PCR analysis of marker genes for NPC populations (*Nestin*, *Sox1*, *Pax6* and *BLBP*) and for differentiated cell types [*βIII-tubulin* (βIII-tub), *GFAP* and *CNPase*]. (**C–F**) Non-purified transduced cells were assayed, and all values given as percentage of transduced cells (i.e. *EGFP* expressing cells; EGFP+ve). (C–E) Dissociated cells were plated onto poly-l-lysine substrate and cultured in the presence of EGF. (C) Quantification of percentage of transduced cells expressing *Pax6* by immunofluorescence. (D) Quantification of percentage of transduced cells expressing activated caspase3 by immunofluorescence. (E) Quantification of transduced cells expressing phospho-histone3 (PH3) by immunofluorescence. (F) Cells were plated onto poly-l-lysine substrate and cultured in the absence of EGF. Percentage of cell types present in differentiated cultures identified by immunofluorescence: progenitor cells (Pax6) and differentiated cells (astrocytes: GFAP; neurons: βIII-tubulin and oligodendrocytes: CNPase). *$P < 0.05$ by Student's *t*-test.

and drives their fate choices towards producing daughter NPCs at the expense of differentiated cell types. As in our previous neuronal cell assays, this result is in general opposite to previous results observed when *HCFC1* is over-expressed in these same cell types and suggests that NPCs are sensitive to *HCFC1* dosage.

Identification of novel HCFC1 variants in individuals with ID

We previously described three missense variants in the *HCFC1* coding region (NM_005334.2) in patients with probable or likely XLID using X-exome sequencing (2). A missense change (c.674G>A, p.Ser225Asn) segregated with ID in family D144, associating with ID in four affected males across two generations (Fig. 3A). A second missense variant (c.2626G>A, p.Gly876Ser) was found in a simplex patient (family D147) whose mother had 100% skewing of X chromosome inactivation. However, this

variant was also found in 1/1466 chromosomes (or 7/10 271; National Heart, Lung, and Blood Institute [NHLBI] Exome Variant Server), and the patient also had a potentially deleterious variant in the XLID gene *MED12* (NM_005120.2; c.3101T>G, p.Phe1034Cys; OMIM: 3 00 188), and a likely deleterious truncating variant in the ID gene *ARID1B* (NM_017519.2; c.2723delC, p.Pro908fs*6; OMIM: 6 14 556) (12) (Fig. 3A). A third *HCFC1* variant (c.5267C>T, p.Ala1756-Val) was found in the index patient of family D82 with four affected males across two generations; however, co-segregation could not be performed as DNA was unavailable on all relevant family members (Fig. 3A). The index patient also had a variant in *ZMYM3* (NM_005096.3; c.356A>G, p.Gln119Arg) of unknown significance. We now report the discovery of two additional novel *HCFC1* variants in patients with probable XLID. The first novel variant was found in a family ascertained in the Czech Republic (family CZE168) that contained two affected brothers. The clinical descriptions of these individuals are detailed in
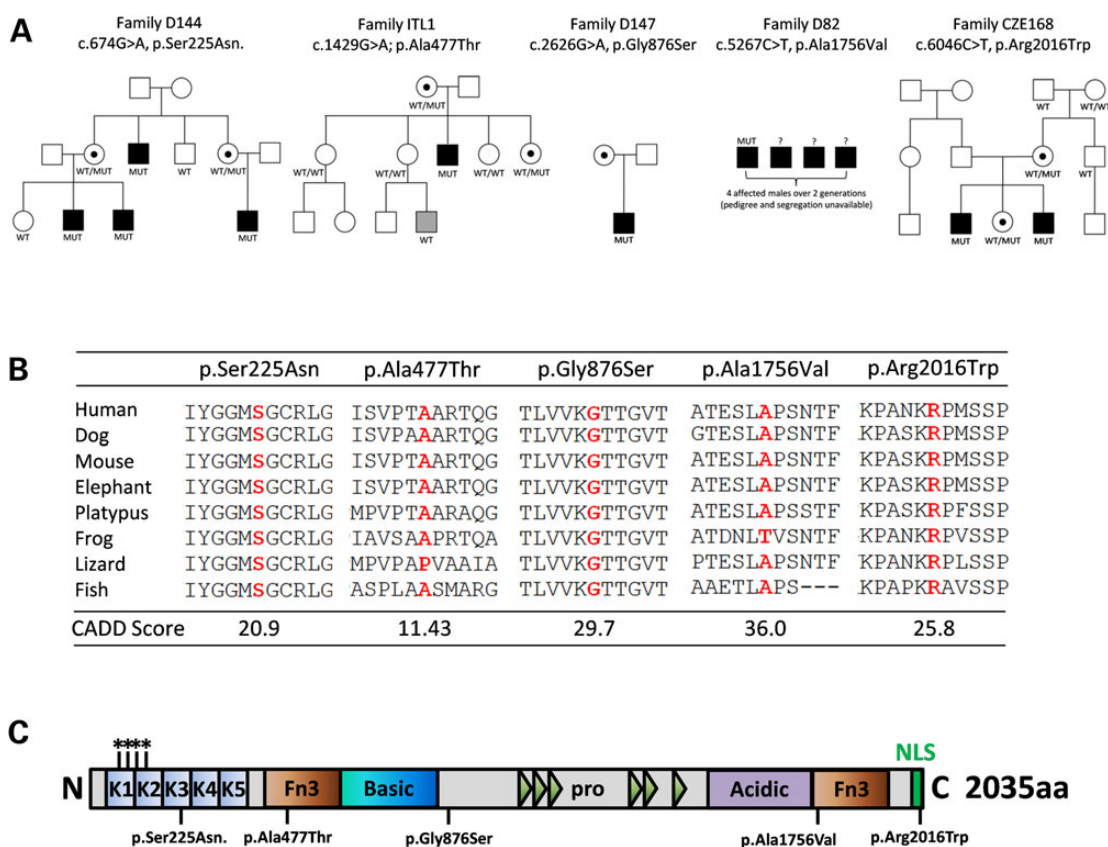
Figure 3. *HCFC1* variants segregate with ID and result in changes to conserved regions of the protein. (**A**) Pedigrees of families identified in (2) (Families D144, D147 and D87) and that of new families (ITL1 and CZE168), with *HCFC1* variants. cDNA (NM_005334.2) and protein (XP_005274721.1) annotation for individual variants are shown. Black fill indicates ID, black dots indicate heterozygote carriers, and grey fill indicates autism (not ID). (**B**) Sequence alignment of variant amino acids across different species (using EBI CLUSTALW server), and *in silico* prediction (CADD) of the effect that the variant amino acid changes have on protein function and integrity. Note that CADD scores >20 are considered likely pathogenic, while scores >10 are considered potentially pathogenic. (**C**) Linear scale cartoon of HCFC1 protein showing known structural features (K1-5: Kelch domains; Fn3: Fibronectin 3 domain; Pro: HCFC1 proteolytic repeats; NLS: nuclear localization sequence) and the locations of variant amino acids. Note that * indicates locations of variants previously reported (9,10).

Supplementary Material, Table S1. Briefly, both affected individuals shared clinical features of mild–moderate ID, delayed speech and delayed psychomotor development. They were both of short stature (3rd percentile) and shared dysmorphic features of thick lips, long nose and a long philtrum. Both displayed mild liver steatosis and diffuse hypocontractibility of the left ventricle. Metabolic profiling also revealed transiently increased plasma homocysteine levels and normal levels of methylmalonic acid in urine. Whole-exome sequencing was performed to find causative variants. No variants were found under an autosomal recessive model, whereas five variants were found under a model of X-linkage (Supplementary Material, Table S2). Of these, the *HCFC1* variant (chrX 153 215 026G>A HG19; c.6046C>T; pArg2016Trp) was the most likely candidate based on extended segregation analysis and predicted tolerances to variation (Supplementary Material, Table S2). X-inactivation studies on the affected individual's mother showed significant skewing (87%). The *HCFC1* variant abolishes a predicted bi-partite nuclear localization sequence (NLS; amino acids 2011–2034; cNLS Mapper (13) prediction score = 9.6). A second novel *HCFC1* variant was found in a family ascertained in Italy (family ITL1; chrX 153 225 268 C>T (HG19); c.1429G>A; p.Ala477Thr; Fig. 3A and Supplementary Material, Fig. S5). The affected individual has severe ID, absent speech and displayed frequent febrile seizures during infancy. He had an elongated face, large ears, arachnodactyly and a lean body habitus. Metabolic profiling revealed normal methylmalonioc levels in

urine. The affected individual's great uncle also suffered from ID (deceased; Supplementary Material, Fig. S5), while the individual's nephew suffered from autism spectrum disorder and did not carry the variant. This *HCFC1* variant disrupts the fibronectin type 3 domain of HCFC1 which is involved in heterodimerization of HCFC1 N- and C-terminal proteolytic cleavage products and DNA binding.

We were able to access material (whole blood) from individuals of the D144, ITL1 and CZE168 families (but not the others) to test for the expression levels of *HCFC1*. Intriguingly, mRNA expression of the variant forms of *HCFC1* in affected individuals was elevated in families ITL1 and CZE168, and modestly elevated in two of three individuals from D144, compared with control samples (Supplementary Material, Fig. S6A–C), which may suggest an autoregulatory mechanism controlling *HCFC1* transcription (see discussion). Western blot analysis of an affected individual in the CZE168 family, however, did not reveal significant increase in HCFC1 protein levels (Supplementary Material, Fig. S6D).

Overall, this new genetic evidence together with previous reports suggests the involvement of the *HCFC1* in the pathology of ID; however, while the genetic evidence for the p.Ser225Asn variant (located within the N-terminal Kelch domain) and p.Arg2016Trp (located in the NLS) is more persuasive, whether the p.Gly876Ser, p.Ala1756Val and p.Ala477Thr variants are likely causative is less clear. All of the variants alter conserved amino acids of HCFC1, and *in silico* predictions (CADD score (14))

suggested four of the variants to be likely deleterious to protein function, with the p.Ala477Thr variant predicted to be possibly deleterious (Fig. 3B and C). We thus embarked on functional characterization of these variants and their effect on HCFC1 function.

The p.Arg2016Trp variant of HCFC1 disrupts its nuclear localization

We cloned four of the five *HCFC1* variants into a cDNA construct that permits expression in mammalian cells under a constitutive promoter. Unfortunately, we were unable to obtain clones with the p.Ala477Thr variant. The expression constructs contained in frame Myc and HA tags at the N and C terminus, respectively, to facilitate discrimination between endogenous and exogenous expression (Fig. 4A). When expressed in HEK293T cells, we observed no difference in *HCFC1* variant mRNA or protein expression levels compared with wild-type *HCFC1*, suggesting the variants did not overtly affect exogenously expressed *HCFC1* mRNA or protein stability (Supplementary Material, Fig. S7A

and B). We next studied the subcellular localization of wild-type and variant HCFC1 proteins. First, we compared the expression of endogenous HCFC1 with that of exogenously expressed wild-type HCFC1. HCFC1 is known to undergo proteolytic processing in cells and can be found as either a full-length protein, or N- and C-terminal fragments (15,16). These fragments are known to form a heterodimer but can also display independent functions (17–19). We used our anti-Myc and HA antibodies to detect exogenous HCFC1, and our anti-HCFC1 antibody, which binds the C-terminal of HCFC1, to localize exogenous and endogenous HCFC1. In untransfected cells, endogenous HCFC1 was localized predominantly in cell nuclei, with highest expression found in cells undergoing mitosis (Fig. 4B). Exogenous HCFC1 detected with the C-terminal Myc tag displayed similar expression, while the HCFC1 detected with the N-terminal HA tag was enriched in the nucleus but also localized to the cytoplasm (Fig. 4B). All variants except the p.Arg2016Trp variant localized akin to wild type. The p.Arg2016Trp variant detected through both its N-terminal HA tag and C-terminal Myc tag failed
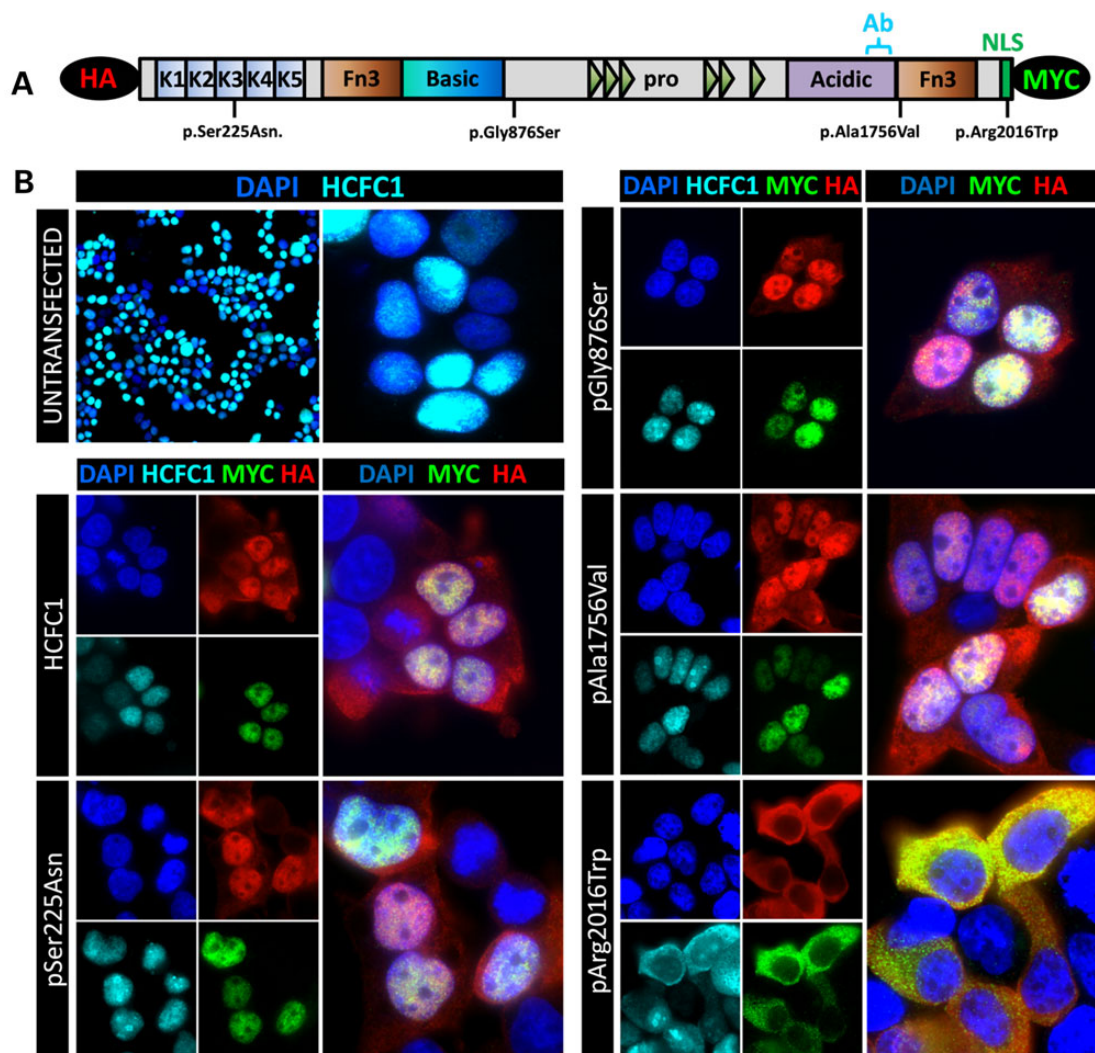
Figure 4. *HCFC1* p.Arg2016Trp variant fails to localize to the nucleus in HEK293T cells. (**A**) Diagram of HCFC1 protein with the locations of the variants being tested. Note also that exogenously expressed HCFC1 proteins are tagged by HA at the N-terminus and Myc at the C-terminus, and also the C-terminal location of the epitope used to generate the HCFC1 antibody (Ab). (**B**) HEK293T cells were transfected with empty control vectors or vectors encoding wild-type or variant forms of *HCFC1*. Immunofluorescent detection of endogenous (and exogenous) HCFC1 (using the anti-HCFC1 antibody; cyan), and exogenous wild-type and variant forms using both anti-HA (red) and anti-Myc (green) antibodies. Cell nuclei are counterstained with DAPI (blue).
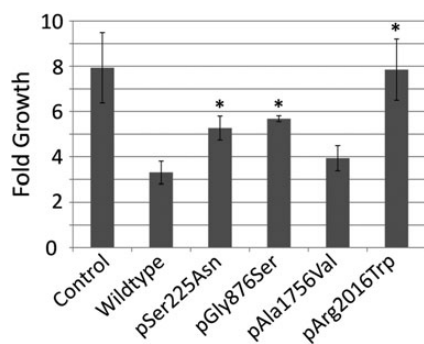
Figure 5. *HCFC1* variants disrupt its ability to suppress cell growth of HEK293T cells. HEK293T cells were transfected with an empty expression vector (control) or with expression vectors encoding either wild-type or variant HCFC1 forms. Transfected cells were plated at equivalent densities and allowed to grow for 2 days at which point cell numbers were assayed. $n = 6$, * is significantly different to wild type, $P < 0.05$ Student's t-test.

Figure 6. Over-expression of *HCFC1* variants disrupts *MMACHC* mRNA expression. HEK293T cells were transfected with an empty expression vector (control) or with expression vectors encoding either wild-type or variant HCFC1 forms and RNA was isolated. *MMACHC* mRNA expression was analysed by qRT–PCR.
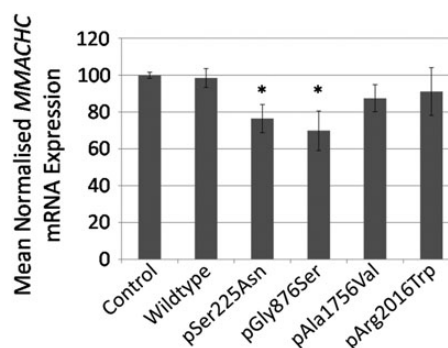
to localize to the nucleus and was found in the cytoplasm. Similar results were obtained using HeLa cells (data not shown) and in hippocampal neurons (Supplementary Material, Fig. S8A and B). These data indicate that the *HCFC1* variant affecting the nuclear localization sequence (p.Arg2016Trp) results in a disruption to its nuclear localization.

HCFC1 variants disrupt proliferation of HEK293T cells

Given previous studies showing the requirement of HCFC1 at multiple stages of the cell cycle (4,6,17–19), we also tested the ability of *HCFC1* and its variant forms to affect the proliferation of HEK293T cells. We plated transfected cells out at an equivalent density and monitored cell growth using the aforementioned cell growth assay. Following 2 days of growth, over-expression of wild-type *HCFC1* caused a 58% decrease in growth compared with control cultures (Fig. 5 and Supplementary Material, Fig. S9). This growth reduction was not significantly different from that resulting from over-expression of p.Ala1756Val (50% decrease). Over-expression of the p.Ser225Asn and p.Gly876Ser mutants both resulted in decreases in growth of 33 and 28%, respectively, which was a modest, but significant loss of growth suppression compared with over-expression of the wild-type *HCFC1*. The over-expression of the p.Arg2016Trp had no effect on growth and was akin to control cells (Fig. 5 and Supplementary Material, Fig. S9). These data suggest that three of four of the tested *HCFC1* variants may cause loss of function to various degrees, with the exception being the p.Ala1756Val variant.

HCFC1 variants disrupt expression of MMACHC

The investigations on loss-of-function *HCFC1* mutations in CblX revealed that *HCFC1* is required for the expression *MMACHC*, which encodes an enzyme involved in the metabolism of cobalamin, and which is also the most common genetic factor found mutated in CblC (9). We thus looked at the ability of the *HCFC1* variants to affect the expression of *MMACHC*. We again utilized our HEK293T cells and over-expressed wild type or variant *HCFC1* and analysed *MMACHC* mRNA expression by qRT–PCR. Over-expression of WT *HCFC1* had no effect on *MMACHC* expression, suggesting that endogenously expressed wild-type HCFC1 binding (and function) is already saturated at the *MMACHC* promoter under normal conditions (Fig. 6). Likewise, expression of the p.Ala1756Val and p.Arg2016Trp had no effect. The other

two variants tested resulted in the reduction of *MMACHC* expression. As we were able to achieve large over-expression of wild-type and variant *HCFC1* in these cells (~160-fold as tested by qRT–PCR, and ~10-fold as tested by western blot; Supplementary Material, Fig. S7), the stoichiometry suggests that the exogenously expressed proteins will outcompete endogenous wild-type HCFC1 for binding at the *MMACHC* promoter. Given that HCFC1 loss-of-function mutations are described to result in reduced *MMACHC* expression in patient cells (9), the loss of *MMACHC* expression in this assay can be interpreted likewise, suggesting that the p.Ser225Asn and p.Gly876Ser variants are also likely loss of function.

HCFC1 variants affect neuronal outgrowth

The above data suggest that with the exception of the p.Ala1756-Val variant, the other HCFC1 variants may cause a loss or partial loss of function. To further investigate this in a model system more relevant to patient phenotypes, we assessed the ability of the four variants to effect the axonal growth of hippocampal neurons. As we have previously demonstrated and as shown above, this assay is both sensitive to *Hcfc1* gain and loss of function, which results in either reduction or extension of axonal growth, respectively (2). We assayed the effect that over-expression of the *HCFC1* variants has on axonal growth and their ability to rescue defects associated with loss of *Hcfc1* expression. We isolated primary hippocampal neurons from E18.5 embryos (as described above) and immediately following isolation, we nucleofected them with an expression vectors encoding Red Fluorescent Protein (RFP) together with either an empty expression vector control or expression vectors encoding either wild-type *HCFC1,* or one of the four *HCFC1* variants. Cells were plated out and allowed to attach overnight. The following day, these transfected cultures were transduced with lentiviral particles that encode both *EGFP* and either control or *Hcfc1*[shRNA]. Cells were grown for an additional 3 days *in vitro*, fixed and immunofluorescently stained for the axonal marker protein TAU1. Axon lengths of dual-transfected and transduced cells were measured (i.e. those expressing both RFP and *EGFP*) (Fig. 7A). Over-expression of *HCFC1* in control-transduced cells resulted in a 25% reduction of axonal length, in agreement with our original findings (2). A similar 23% reduction was observed when the p.Ala1756Val variant was over-expressed, while the effect of over-expression of the other three variants resulted in less severe axonal length reductions (ranging only between 4 and 8%) that were significantly different from
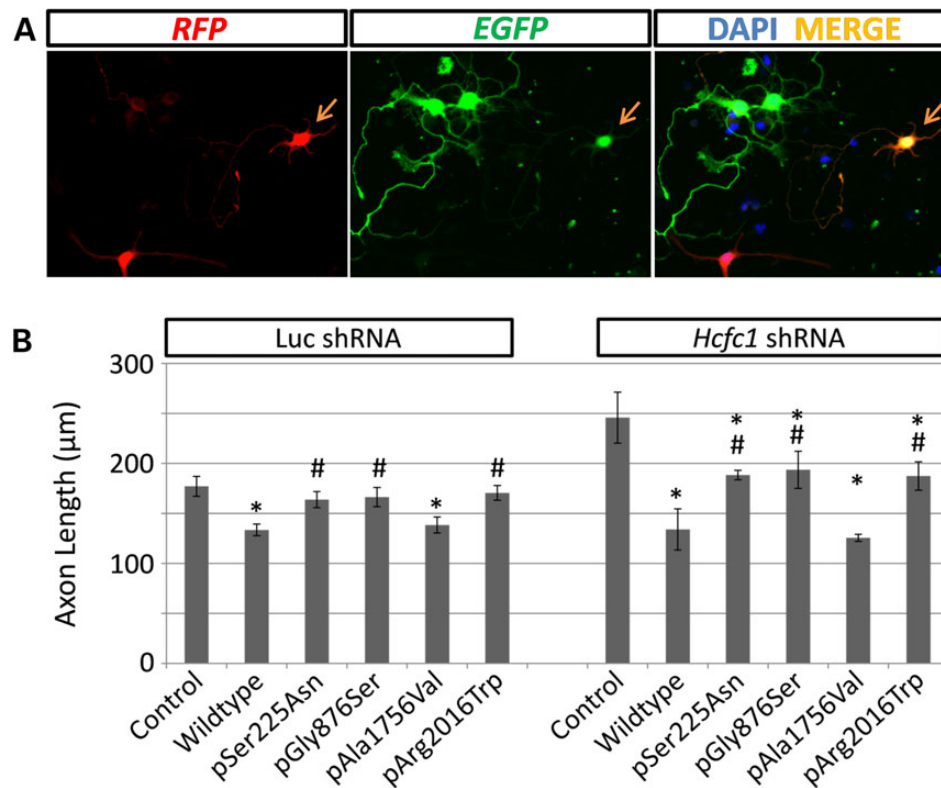
Figure 7. *HCFC1* variants disrupt axon growth. Hippocampal neurons were isolated and transfected with *RFP* expression vector together with either an empty expression vector (control) or vectors encoding either wild-type or variant forms of *HCFC1*. Following cell attachment *in vitro*, neurons were transduced with lentiviral particles that encode *EGFP* and either control or *Hcfc1*[shRNA]. (**A**) Representative immunofluorescent image from control condition showing identification of a dual transfected (i.e. RFP expressing; red) and transduced (i.e. EGFP expressing; green) neuron (yellow arrow). Nuclei were stained using DAPI (blue). (**B**) Quantification of mean axonal length. $P < 0.05$ by Student's *t*-test comparing values within a shRNA condition (Luc or *Hcfc1*[shRNA]); *significantly different to control condition; #significantly different to wild-type condition.

HCFC1 wild-type over-expression. Next, we analysed samples transduced with *Hcfc1*[shRNA] and revealed similar outcomes (Fig. 7B). Transduction with this virus caused a 39% increase in axonal length compared with transduction with a control virus. This elongated axon length was reduced by 45% when *HCFC1* wild-type was re-expressed and a similar 49% reduction observed when the p.Ala1756Trp variant was expressed. Re-expression of the other three *HCFC1* variants caused significant reductions (i.e. compared with control *Hcfc1*[shRNA] condition) that ranged between 21 and 24%; however, these reductions were significantly lower, compared with re-expression of the wild-type *HCFC1* (Fig. 7B). Together, these results confirm that axonal length is regulated by *HCFC1* dosage and that three of the four *HCFC1* variants tested here (i.e. p.Ser225Asn, p.Gly876Ser and p.Arg2016Trp) lead to partial HCFC1 loss of function.

Discussion

Previous findings suggest that both HCFC1 gain and loss of function mutations can cause neurodevelopmental disorders. Here, we complement our previous studies on *HCFC1* gain of function by reporting the discovery of altered neural cell behaviours resulting from *HCFC1* loss of function. We find, in general opposite to the effects of over-expression, that modest knockdown of *Hcfc1* in NPCs caused over-proliferation and an associated reduction of differentiation, while knockdown in hippocampal neurons promoted axon growth. In support of these findings, brain malformations including microcephaly, macrocephaly and cortical

gyration malformation have been reported in CblX patients (9,10) and in individuals from families CZE168 and D147 that we describe here. Thus, embryonic brain growth and neuronal connectivity are processes sensitive to HCFC1 function (and dosage), and alterations to these processes may underlie aspects of the patient neurological phenotypes.

Having established neural-specific effects of loss of *Hcfc1* function, we sought to provide evidence in support of the pathogenicity of additional *HCFC1* variants described previously (2), and newly discovered, including a novel variant affecting the NLS of HCFC1. Three of the four variants we tested resulted in altered HCFC1 function consistent with a partial, but not complete, loss-of-function effect with the exception being the p.Ala1756Val variant. The other three variants (p.Ser225Asn, p.Gly876Ser and p.Arg2016Trp) all showed reduced (but not abolished) potency compared with wild type when assayed for effects on HEK293T proliferation and neuronal axonal growth. A similar effect was observed for the p.Ser225Asn and p.Gly876Ser variants when assayed for an ability to regulate *MMACHC* transcription levels, but not the p.Arg2016Trp variant, which behaved akin to wild type (and p.Ala1756Val variant) in this instance. The inability of the HCFC1 p.Arg2016Trp variant to effect *MMACHC* expression in a negative manner is however consistent with its exclusion from the nucleus, which would mask its potential ability to disrupt transcription carried out by endogenous wild-type HCFC1 in these cells. In fact, by measure of the effect on HEK293T proliferation and localization, the p.Arg2016Trp variant may be considered the most detrimental to protein function of the four

variants tested. Furthermore, the two affected brothers with this variant both displayed elevated homocysteine levels, a signature metabolic feature of individuals with CblX (and CblC), and thus suggestive of loss of HCFC1 transcriptional activation of *MMACHC*. Interestingly, *HCFC1* mRNA levels were found to be elevated in whole blood of the affected individuals and also in the individual with the p.Ala477Thr mutation, family ITL1, and in lymphoblastoid cell lines of two of the three individuals with the p.Ser225Asn mutation, family D144. This finding may suggest that HCFC1 normally participates in the repression of its own transcription. Such an autoregulatory mechanism would be in line with the need to tightly regulate dosage of HCFC1 during development and homeostasis. That HCFC1 binds YY1, a known repressor of *HCFC1* transcription, provides a candidate mechanism warranting further investigation, particularly in light of our previous finding wherein mutation of a YY1 binding site in the *HCFC1* 5′UTR was associated with *HCFC1* over-expression and ID [2,20]. Although transcription was elevated, initial investigations did not show an increase in the overall protein level of the p.Arg2016Trp variant.

Despite the deleterious effect on HCFC1 function in our cell assays, the contribution of the p.Gly876Ser variant to the pathology of the patient remains complicated by the likely pathogenic contributions of a *MED12* variant (c.3101T>G, p.Phe1034Cys) and a protein truncating *ARID1B* variant (c.2723delC, p.Pro908fs*6) also found in the individual. Mutations in *MED12* cause XLID [21] (OMIM: 3 00 895 and 3 09 520). This particular variant has not been previously reported and is predicted to be potentially pathogenic (CADD score = 14.53); as such its role in this individual's phenotype is not conclusive. In contrast, haploinsufficiency of *ARID1B* invariably causes a broad range of disorders from mild ID to more severe phenotypes of Coffin-Siris syndrome (OMIM: 1 35 900) and can be associated with a spectrum of clinical presentations [12,22]. Because of this, it is difficult to disentangle the contributions of *HCFC1*, *MED12* or *ARID1B* variants. Unfortunately, we were unable to obtain information on the affected individual's homocysteine or methylmalonic acid levels, which could potentially identify CblX. The aggregate clinical data available to us are still consistent with CblX, but also severe *ARID1B*-associated phenotypes, and have some features not yet described for *MED12* mutations (e.g. epilepsy), and missing some features normally associated with MED12 mutations (e.g. hypertelorism and constipation). It is credible that *HCFC1*, *ARID1B* and *MED12* variants may all be contributing to the ID; the possibility of multiple hits with additive or independent effect has been previously documented and might be under-ascertained [23,24]. Interestingly, the EVS server reports the p.Gly876Ser variant in 7/10 271 chromosomes (including two males with the change, a male frequency of 1/1193). Furthermore, following completion of our functional studies, the ExAC database became available (http://exac.broadinstitute.org, last accessed February 2015), which also reported the p.Gly876Ser variant at a frequency of 52/1 21 298 chromosomes (of which 17 are likely to be males, giving an approximate male frequency of ~1/2300). These data reveal that the p.Gly876Ser occurs at a very low frequency in the population and suggest that the variant might be benign; however, it is also possible that the males with this variant in these databases may have mild to borderline ID, which aligned with the high incidence (affecting 2–3%) and wide spectrum of ID is plausible, albeit undeterminable at this point in time. Our conservative conclusion based on the functional and genetic data presented here is to regard the p.Gly876Ser variant as a low-frequency variant that provides an increased susceptibility to intellectual disability. This is consistent with recent evidence which reveals that complex

disease traits are associated with collections of common variants found in monogenic disease genes [25], and that milder neurodevelopmental phenotypes (for example autism spectrum disorder, and probably mild ID) are more likely to involve combinations of rare and common variants [26,27]. That the p.Gly876Ser variant has been identified in a patient with autism may provide additional support of its involvement [28].

Our data indicate that of the four variants we tested, three affect HCFC1 protein function, namely the p.Ser225Asn, p.Gly876-Ser and p.Arg2016Trp variants found in families D144, D147 and CZE168, respectively. The p.Arg2016Trp variant disrupted a nuclear localization signal sequence in the C-terminus of the protein. Our localization studies show that this indeed did disrupt the ability of HCFC1 to become enriched in the nucleus; however, that the protein retained some function suggested that nuclear import was reduced, but not abolished. The p.Ser225Asn variant lies in the Kelch repeat domain of HCFC1 similar to previously reported loss-of-function HCFC1 variants [9]. This domain is characterized as a protein–protein interaction motif and has been shown to mediate binding to transcription factors (for example E2F1 and E2F4 and THAP11) and chromatin regulators (such as MLL H3K4 Methyltransferase) [4,29]. HCFC1 has been described as an obligate co-factor of THAP11, which recruits HCFC1 to target promoters, including that of the *MMACHC* gene [9,29,30]. Thus, the p.Ser225Asn variant may disrupt these interactions, as has been postulated for other variants of HCFC1 affecting the Kelch domain [9]. Likewise, the p.Gly876Ser variant may affect the protein–protein interactions mediated by the basic domain, which includes DNA-binding proteins (e.g. ZBTB17) and chromatin modifiers such as SIN3A [18]. It will be interesting to identify which promoters and genes are affected on a global scale using patient-derived cell lines, to better understand the biology of HCFC1 and understand the genetic pathways that underlie ID. In all of our assays, the p.Ala1756Val variant behaved as wild-type HCFC1. While this does not support its pathogenicity, it does not exclude its involvement in some finer way and may reflect limitations of the assay conditions to detect subtle disruptions to HCFC1 function (see below). Such a small disruption may still be solely responsible for pathogenesis, or alternatively its effect may be compounded by contributions of other rare variants (e.g. the c.356A>G, p.Gln119Arg variant in ZMYM3 of unknown significance discovered in the patient) or common variants [24,25].

As aforementioned, loss-of-function mutations that alter the Kelch domain of HCFC1 have been identified in patients with CblC-like disorder called CblX [9,10]. CblC is most commonly caused by mutations in *MMACHC* and results in methylmalonic aciduria and homocystinuria. The defect causes decreased levels of the coenzymes adenosylcobalamin (AdoCbl) and methylcobalamin (MeCbl), which results in decreased activity of the respective enzymes methylmalonyl-CoA mutase and methyltetrahydrofolate:homocysteine methyltransferase, also known as methionine synthesis. The main biochemical findings in CblC patients with *MMACHC* mutations are increased plasma homocysteine level, low or normal plasma metionin level, homocysteinuria and methylmalonic aciduria [31]. Evidence from CblX patient cell lines and knockdown experiments in HEK293T cells revealed that these *HCFC1* mutations resulted in an almost complete loss of *MMACHC* expression, suggesting almost complete loss of HCFC1 function [9]. The patients from whom the variants we tested here have in general much milder neurological presentations compared with CblX (and CblC) individuals and consistently, in our assays, it appears that the HCFC1 retained at least some biological activity, i.e. our assays suggest only partial loss of function

(Supplementary Material, Table S3). Furthermore, both the individuals with p.Arg2016Trp mutation had transiently elevated homocysteine levels, but repeatedly normal levels of plasma metionin and methylmalonic acid in urine (the measurement of AdoCbl and MeCbl was not available). Normal urine methylmalonic acid levels were also reported for the individual with the p.Ala477Thr variant. These nearly normal biochemical findings correspond to the much milder neurological defects compared with CblC patients, again suggesting partial MMACHC activity, and only a partial defect in HCFC1 function. It remains to be tested to what extent *MMACHC* expression may be affected in these individuals, and the other individuals, and we studied here (materials unavailable at this time), and more so, to what extent the metabolism of cobalamin is affected in all. Such information may have some clinical value, as expression levels of mutant *MMACHC* alleles have been associated with age of onset and severity in CblC patients (11). However, as CblC causing *MMACHC* mutations are all reported as homozygous (or compound heterozygous), it suggests that unless MMACHC function is severely compromised, the metabolic symptoms may not exist (11). That the metabolic features of CblX disorder are relatively mild compared with CblC patients, but the neurological phenotype is much more severe (includes brain malformation and intractable epilepsies) suggests that *MMACHC* deficiency alone does not explain the neurological phenotype, and that *HCFC1* is important for the expression of additional genes important for normal brain development (9). Thus, mutations resulting in partial loss of HCFC1 function may result in neurological phenotypes in the absence of the clinical presentation of cobalamin deficiency. Furthermore, while treatment of early-onset CblC patients with a combinatorial approach consisting of supplementation of hydroxycobalamin, betain and folic acid provides some improvements in the visceral and haematological symptoms, the efficacy on neurological outcomes is not established, and patients invariably continue to show neurological and cognitive impairment regardless of time of diagnosis and/or treatment initiation (32).

The mounting challenge of assigning pathogenicity to large numbers of variants is typically approached by employing genetic and statistical evidence often without in-depth functional studies or even detailed clinical assessment. By comprehensive investigations of four *HCFC1* variants using all above criteria, we found conforming evidence for pathogenicity of p.Ser225Asn and p.Arg2016Trp; however, some discordance was observed with the remaining two variants studied. The variant with the highest CADD score, p.Ala1756Val, which is also not present in any publicly available control dataset we analysed, behaved as wild type in our cell-based studies. Conversely, the p.Gly876Ser variant which is found at very low frequency in northern European descent population displayed loss of function in our cell assays. This raises important issues pertaining to the relevance of such functional tests for reaching correct clinical diagnosis and actions thereafter. Resolution of these fundamental questions, also relevant to many other genes and disorders, will involve deeper understanding of the function of these genes, better genotype–phenotype databases, more sophisticated models (e.g. patient-derived stem cells) or simply more patients and variants investigated. Our work underlines the importance of combined approach to the resolution of the relevance of genetic variation in neurological disease, involving clinical, genetic, bioinformatics and cell and molecular functional studies.

In conclusion, our data provide additional support for the involvement of *HCFC1* in ID. Together, with previous studies, it is now apparent that both gain and loss of HCFC1 function can result in changes to the behaviour of embryonic neural cells, which

likely underpins aspects of patient phenotypes (2,33). Regarding loss-of-function mutations, the severity of the phenotypes likely reflects the extent to which the mutations affect HCFC1 function, with almost complete loss of function causing CblX featuring severe neurological phenotypes (e.g. retractable epilepsy, brain malformations, severe neurocognitive impairment) and clinically relevant deficiencies in cobalamin, whereas partial loss of function resulting in less severe neurological phenotypes (e.g. mild ID) (Supplementary Material, Table S3) (9). The continued emergence of large-scale sequencing practices in research and diagnostics will likely provide further insight into the spectrum of phenotypic outcomes stemming from both gain- and loss-of-function mutations in *HCFC1*.

Materials and Methods

Animal use

This study was performed under regulations of the South Australian Animal Welfare Act 1986, and in strict accordance with the Australian Code of Practice for the Care of Animals for Scientific Purposes, 2004. The protocol was approved by the Women's and Children's Health Network (WCHN) Animal Ethics Committee (Approval Number: 944/03/15). All euthanasia was performed using cervical dislocation, and every effort was made to minimize suffering. Time-mated pregnant female Swiss mice were obtained from the Women's and Children's Health Network Animal Care Facility (Women's and Children's Hospital, Adelaide, Australia).

Generation of lentivirus

Constitutive 3rd generation lentiviral vectors were employed as previously described (34–36). To generate *Hcfc1*-specific shRNA transfer vectors, three shRNA sequences predicted to specifically target *Hcfc1* UTRs were designed using BLOCK-it RNAi designer software [Invitrogen, Carlsbad, CA, USA)]. A shRNA sequence targeted to luciferase was used as a control (36). The shRNA sequences are T1: GCAGAAGGCAGATTGGAAAGA; T2: GGAGATAACACCCA TACTTAA and T3: GCACTTGTT TGTAGT TCATCG. These sequences were cloned into the lentiviral transfer vector plv-C-shRNA, and lentiviral particle stocks were generated and titrated as previously described (35,36). Transduction of neurons was performed at either multiplicity of infection (MOI) = 1 overnight (O/N) at Day 1 of culture. Transduction of NPCs was achieved at MOI = 20 using O/N incubation with dissociated cells derived from secondary neurospheres. For purification of transduced cells, neurospheres were dissociated and EGFP expressing cells purified using FACS (BD FACSAriaII flow cytometer; BD Bioscience, San Jose, CA, USA), for further culture as neurospheres.

Cloning of HCFC1 expression constructs

pCGN-HCFC1-FL vector was a kind gift from W. Herr, Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland (15). To generate mutant constructs, fragments containing the desired nucleotide(s) for mutagenesis were first amplified and then subcloned into pGEMT by pGEM®-T Easy Vector Systems (Promega, Madison, WI, USA) using the following primer pairs: for p.Ser225Asn, F-TATGACGTGCCTGACTATGC and R-TAGG TACCAGAGGTCCTTGC; for p.GlySer876, F-AGTAGCCCACAGAT GAGTGG and R-CTTGAGAGGGTGGCAATGG; for p.Ala1756Val and p.Ala2016Trp, F-TTATCACTTCCCCAAGAGC and R-CTCACCCT GAAGTTCTCAGG. Mutagenesis reaction was performed directly on pGEMT vector using QuikChange Multi Site-Directed

Mutagenesis Kit (Stratagene, La Jolla, CA, USA) according to the manufacturer's protocol using the following primers: for p.Ser225Asn, F-TACGGCGGGATGAATGGCTGCAGGCTGG; for p.GlySer876, F-CGTTGGTTGTGAAAAGCACCACAGGTGTCACG; for p.Ala1756Val, F-CACTGAGAGCCTGGTTCCATCCAACACATTTGTGG; and for p.Ala2016Trp, F-CCAAGCCAGCCAACAAGTGGCCCATGTCCTCTC. Subsequently, mutated fragments were subcloned back into pCGN-HCFC1-FL construct by restriction digest followed by ligation with T4 ligase (NEB, Ipswich, MA, USA): for p.Ser225Asn, *SexAI* and *XbaI* ; for p.GlySer876, *NotI* and *SpeI*; for p.Ala1756Val or p.Ala2016Trp, *BamHI* and *SgrA1*. All cloning steps were performed using DH5α bacterial strain. Vector extraction was performed using Miniprep Kit (QIAGEN, Germantown, MD, USA) according to the manufacturer's protocol. For transfection and virus production, vectors were purified using Endofree Plasmid Maxi Kit (QIAGEN).

Cell culture

Isolation of NPCs from the E18 cortex was as previously described (36–38). Single cells for NPC assays were generated by passing dissociated cells through a 0.75 µm cell filter (BD Biosciences). All neural progenitor cell assays were conducted in biological triplicate, with each replicate representing a separate transduction experiment. Averages of these triplicate results are reported. The cell proliferation assay was conducted using the Cell Titre 96AQueous Kit (Promega) as per manufacturer's instruction; each biological replicate was analysed using technical quadruplicates and normalized against starting cell numbers (±4.2% cells per well). Adherent NPC assays were performed by plating dissociated neurosphere cells onto poly-l-lysine (Sigma-Aldridge, St Louis, MO, USA) coated coverslips (Menzel-glasser, Thermo Fisher Scientific, Waltham, MA, USA) at $1 \times 10^4/cm^2$, respectively, and culturing cells in neurosphere media with or without EGF. Immunofluorescent staining and fluorescent microscopy (see below) were used for cell count analysis. For each cell count analysis, at least 200 cells were scored for each replicate experiment (+EGF condition: pax6 total $n = 1230$ control, and $n = 1486$ $Hcfc1^{shRNA}$; phospho-Histone3: total $n = 1168$ control, and $n = 1208$ $Hcfc1^{shRNA}$; activated caspase3: total $n = 602$ control, and $n = 613$ $Hcfc1^{shRNA}$; and –EGF condition: CNPase: total $n = 780$ control, and $n = 602$ $Hcfc1^{shRNA}$; GFAP: total $n = 1160$ control; $n = 1130$ $Hcfc1^{shRNA}$; βIII-tubulin: total $n = 640$ control, and $n = 603$ $Hcfc1^{shRNA}$; Pax6: total $n = 680$ control; $n = 1190$ $Hcfc1^{shRNA}$). Isolation of primary hippocampal neurons was as described previously (2). Nucleofection was conducted as previously described (2) using 5 µg of pCGN-HCFC1 expression vectors (wild type or variant) and 1 µg of pDsRed-Monomer-C1 (RFP expression vector; Clontech, Mountain View, CA, USA). Morphometric analysis was conducted using immunofluorescent staining and fluorescent microscopy (see below). All data represent average of three biological replicates. For each replicate, at least 30 neurons were scored using the ImageJ software package (NIH). HEK293T and NIH3T3 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum (Gibco, Life Technologies, Mulgrave, Victoria, Australia). Transfections were conducted using Lipofectamine 3000 as per manufacturer's instruction (Invitrogen).

Immunofluorescence

Cultured cells were fixed with 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) for 15 min at room temperature (R/T). Cells were block permeabilized using PBS containing 0.1% Tween20 (PBST) with either 10% normal horse serum (NHS) or

5% bovine serum albumin (BSA) for 1 h at R/T. Primary and secondary antibodies were incubated in PBST with either 3% NHS or 1% BSA overnight (O/N) at 4°C and 1 h at RT, respectively, at the following dilutions; chicken anti-MAP2, mouse anti-CNPAse, mouse anti-Tau1 (all at 1:2000; Chemicon, Millipore Bioscience Research Reagents, Temecula, CA, USA), rabbit anti-GFAP, mouse anti-βIII-tubulin, (both at 1:300; Sigma-Aldridge), rabbit anti-Pax6 (1:200; Chemicon), rabbit anti-phospho-histone3 (1:200, Abcam, Cambridge, UK) Rabbit anti-activated caspase3 (1:200, BD Biosciences), mouse anti-Myc (1:500, Santa Cruz, Santa Cruz Biotechnology, Santa Cruz, CA, USA), chicken anti-HA (Abcam, 1:500), rabbit anti-HCFC1 (1:200, Bethyl Laboratories, Montgomery, TX, USA), donkey anti-sheep Alexafluor555, donkey anti-rabbit Alexafluor488/555/647 and donkey anti-mouse Alexafluor488/555/647 (all 1:1000; Invitrogen), donkey anti-chickenCy3 (Jackson Laboratories, Bar Harbour, ME, USA). Cells were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) and mounted with Slow-fade mounting media (both from Invitrogen). Non-specific staining was controlled by using secondary-only controls (data not shown). Fluorescence was viewed using the Axioplan2 microscope (Carl Zeiss, Jena, Germany) fitted with an HBO 100 lamp (Carl Zeiss). Images were captured using an Axiocam Mrm camera and Axio Vs40 v4.5.0.0 software (Axiovision, Carl Zeiss).

Biochemical analysis

Protein was isolated from cells using a lysis buffer (120 mM NaCl, 50 mM Tris–HCl (pH 8.0), 0.5% NP-40 (v/v), 1 × protease inhibitor cocktail (Sigma), 1 mM Na_3VO_4, 1 mM NaF, 1 mM PMSF) and quantitated using Bradford assay (Biorad, Hercules, CA, USA). Samples were separated and transferred to nitrocellulose (Biotrace NT, Pall Corporation, New York, NY, USA) using the NuPage precast gel system as per manufacturer's instructions (Invitrogen). Blots were blocked using PBST containing 5% NHS and 5% skim milk. Antibodies were incubated in this same solution at the following dilutions; rabbit anti-HCFC1 (1:2000), mouse anti-β-actin (1:5000; Sigma-Aldridge), goat anti-rabbit-HRP and goat anti-mouse HRP (both 1:1000; Dako, Glostrup, Denmark). RNA was isolated from cells/tissues using TRIzol reagent as per manufacturer's instructions (Invitrogen) and further processed using the RNAeasy kit including DNAse treatment (Qiagen). cDNA was generated using SuperscriptIII reverse transcriptase as per manufacturer's instructions (Invitrogen) using random hexamer priming (Geneworks, Adelaide, Australia). The StepOne platform and software (Applied Biosciences, Invitrogen) was employed for qRT-PCR analysis. PCR reactions were generated using the sybr green master mix as per manufacturer's instructions (Biorad) and run using the following parameters: 95°C—5 min; 35 cycles of: 95°C—10 s, 60°C—30 s; followed a melt curve increment step 60°C–100°C. All primers used met strict quality control parameters amplifying control cDNA at 100 ± 10% efficiencies and producing only single PCR products. Primers for *β-Actin*, *HCFC1* and *MMACHC* are as previously described (2,9).

Supplementary material

Supplementary material is available at *HMG* online.

Conflict of Interest statement. None declared.

Funding

References

1. Ropers, H.H. (2008) Genetics of intellectual disability. *Curr. Opin. Genet. Dev.*, **18**, 241–250.
2. Huang, L., Jolly, L.A., Willis-Owen, S., Gardner, A., Kumar, R., Douglas, E., Shoubridge, C., Wieczorek, D., Tzschach, A., Cohen, M. *et al.* (2012) A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability. *Am. J. Hum. Genet.*, **91**, 694–702.
3. Vercauteren, K., Gleyzer, N. and Scarpulla, R.C. (2008) PGC-1-related coactivator complexes with HCF-1 and NRF-2beta in mediating NRF-2(GABP)-dependent respiratory gene expression. *J. Biol. Chem.*, **283**, 12102–12111.
4. Tyagi, S., Chabes, A.L., Wysocka, J. and Herr, W. (2007) E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases. *Mol. Cell*, **27**, 107–119.
5. Julien, E. and Herr, W. (2004) A switch in mitotic histone H4 lysine 20 methylation status is linked to M phase defects upon loss of HCF-1. *Mol. Cell.*, **14**, 713–725.
6. Goto, H., Motomura, S., Wilson, A.C., Freiman, R.N., Nakabeppu, Y., Fukushima, K., Fujishima, M., Herr, W. and Nishimoto, T. (1997) A single-point mutation in HCF causes temperature-sensitive cell-cycle arrest and disrupts VP16 function. *Genes Dev.*, **11**, 726–737.
7. Michaud, J., Praz, V., James Faresse, N., Jnbaptiste, C.K., Tyagi, S., Schutz, F. and Herr, W. (2013) HCFC1 is a common component of active human CpG-island promoters and coincides with ZNF143, THAP11, YY1, and GABP transcription factor occupancy. *Genome Res.*, **23**, 907–916.
8. Hu, H., Haas, S.A., Chelly, J., Van Esch, H., Raynaud, M., de Brouwer, A.P., Weinert, S., Froyen, G., Frints, S.G., Laumonnier, F. *et al.* (2015) X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. *Mol. Psychiatry*, doi:10.1038/mp.2014.193.
9. Yu, H.C., Sloan, J.L., Scharer, G., Brebner, A., Quintana, A.M., Achilly, N.P., Manoli, I., Coughlin, C.R. 2nd, Geiger, E.A., Schneck, U. *et al.* (2013) An X-linked cobalamin disorder caused by mutations in transcriptional coregulator HCFC1. *Am. J. Hum. Genet.*, **93**, 506–514.
10. Gerard, M., Morin, G., Bourillon, A., Colson, C., Mathieu, S., Rabier, D., Billette de Villemeur, T., Ogier de Baulny, H. and Benoist, J.F. (2015) Multiple congenital anomalies in two boys with mutation in HCFC1 and cobalamin disorder. *Eur. J. Med. Genet.*, **58**, 148–153.
11. Lerner-Ellis, J.P., Anastasio, N., Liu, J., Coelho, D., Suormala, T., Stucki, M., Loewy, A.D., Gurd, S., Grundberg, E., Morel, C.F. *et al.* (2009) Spectrum of mutations in MMACHC, allelic expression, and evidence for genotype-phenotype correlations. *Hum. Mutat.*, **30**, 1072–1081.
12. Wieczorek, D., Bogershausen, N., Beleggia, F., Steiner-Haldenstatt, S., Pohl, E., Li, Y., Milz, E., Martin, M., Thiele, H., Altmuller, J. *et al.* (2013) A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.*, **22**, 5121–5135.
13. Kosugi, S., Hasebe, M., Tomita, M. and Yanagawa, H. (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl Acad. Sci. USA*, **106**, 10171–10176.
14. 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. *Nat. Genet.*, **46**, 310–315.
15. Wilson, A.C., LaMarco, K., Peterson, M.G. and Herr, W. (1993) The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein. *Cell*, **74**, 115–125.
16. Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J.W., Conaway, R.C. and Herr, W. (2011) O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. *Cell*, **144**, 376–388.
17. Julien, E. and Herr, W. (2003) Proteolytic processing is necessary to separate and ensure proper cell growth and cytokinesis functions of HCF-1. *EMBO J.*, **22**, 2360–2369.
18. Mangone, M., Myers, M.P. and Herr, W. (2010) Role of the HCF-1 basic region in sustaining cell proliferation. *PLoS ONE*, **5**, e9020.
19. Wilson, A.C., Freiman, R.N., Goto, H., Nishimoto, T. and Herr, W. (1997) VP16 targets an amino-terminal domain of HCF involved in cell cycle progression. *Mol. Cell. Biol.*, **17**, 6139–6146.
20. Yu, H., Mashtalir, N., Daou, S., Hammond-Martel, I., Ross, J., Sui, G., Hart, G.W., Rauscher, F.J. 3rd, Drobetsky, E., Milot, E. *et al.* (2010) The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression. *Mol. Cell. Biol.*, **30**, 5071–5085.
21. Risheg, H., Graham, J.M. Jr., Clark, R.D., Rogers, R.C., Opitz, J.M., Moeschler, J.B., Peiffer, A.P., May, M., Joseph, S.M., Jones, J.R. *et al.* (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat. Genet.*, **39**, 451–453.
22. Hoyer, J., Ekici, A.B., Endele, S., Popp, B., Zweier, C., Wiesener, A., Wohlleber, E., Dufke, A., Rossier, E., Petsch, C. *et al.* (2012) Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.*, **90**, 565–572.
23. Wright, C.F., Fitzgerald, T.W., Jones, W.D., Clayton, S., McRae, J.F., van Kogelenberg, M., King, D.A., Ambridge, K., Barrett, D.M., Bayzetinova, T. *et al.* (2014) Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet*, doi: 10.1016/S0140-6736 (14)61705-0.
24. Girirajan, S. and Eichler, E.E. (2010) Phenotypic variability and genetic susceptibility to genomic disorders. *Hum. Mol. Genet.*, **19**, R176–R187.
25. Blair, D.R., Lyttle, C.S., Mortensen, J.M., Bearden, C.F., Jensen, A.B., Khiabanian, H., Melamed, R., Rabadan, R., Bernstam, E.V., Brunak, S. *et al.* (2013) A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. *Cell*, **155**, 70–80.
26. Gaugler, T., Klei, L., Sanders, S.J., Bodea, C.A., Goldberg, A.P., Lee, A.B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J. *et al.* (2014) Most genetic risk for autism resides with common variation. *Nat. Genet.*, **46**, 881–885.
27. Schaaf, C.P., Sabo, A., Sakai, Y., Crosby, J., Muzny, D., Hawes, A., Lewis, L., Akbar, H., Varghese, R., Boerwinkle, E. *et al.* (2011) Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum. Mol. Genet.*, **20**, 3366–3375.
28. Piton, A., Gauthier, J., Hamdan, F.F., Lafreniere, R.G., Yang, Y., Henrion, E., Laurent, S., Noreau, A., Thibodeau, P., Karemera, L. *et al.* (2011) Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. *Mol. Psychiatry*, **16**, 867–880.

29. Parker, J.B., Palchaudhuri, S., Yin, H., Wei, J. and Chakravarti, D. (2012) A transcriptional regulatory role of the THAP11-HCF-1 complex in colon cancer cell function. *Mol. Cell. Biol.*, **32**, 1654–1670.

30. Dejosez, M., Levine, S.S., Frampton, G.M., Whyte, W.A., Stratton, S.A., Barton, M.C., Gunaratne, P.H., Young, R.A. and Zwaka, T.P. (2010) Ronin/Hcf-1 binds to a hyperconserved enhancer element and regulates genes involved in the growth of embryonic stem cells. *Genes Dev.*, **24**, 1479–1484.

31. Rosenblatt, D.S. and Fowler, B. (2006) In Fenandes, J., Saudubray, J.M., van den Berghe, G. and Walter, J.H. (eds), *Inborn Metabolic Diseases, Diagnosis and Treatment*. Springer, Heidelberg, Germany, pp. 379–396.

32. Martinelli, D., Deodato, F. and Dionisi-Vici, C. (2011) Cobalamin C defect: natural history, pathophysiology, and treatment. *J. Inherit. Metab. Dis.*, **34**, 127–135.

33. Quintana, A.M., Geiger, E.A., Achilly, N., Rosenblatt, D.S., Maclean, K.N., Stabler, S.P., Artinger, K.B., Appel, B. and Shaikh, T.H. (2014) Hcfc1b, a zebrafish ortholog of HCFC1, regulates craniofacial development by modulating MMACHC expression. *Dev. Biol.*, **396**, 94–106.

34. Brown, C.Y., Sadlon, T., Gargett, T., Melville, E., Zhang, R., Drabsch, Y., Ling, M., Strathdee, C.A., Gonda, T.J. and Barry, S.C. (2010) Robust, reversible gene knockdown using a single lentiviral short hairpin RNA vector. *Hum. Gene Ther.*, **21**, 1005–1017.

35. Barry, S.C., Harder, B., Brzezinski, M., Flint, L.Y., Seppen, J. and Osborne, W.R. (2001) Lentivirus vectors encoding both central polypurine tract and posttranscriptional regulatory element provide enhanced transduction and transgene expression. *Hum. Gene Ther.*, **12**, 1103–1108.

36. Jolly, L.A., Homan, C.C., Jacob, R., Barry, S. and Gecz, J. (2013) The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. *Hum. Mol. Genet.*, **22**, 4673–4687.

37. Basak, O. and Taylor, V. (2007) Identification of self-replicating multipotent progenitors in the embryonic nervous system by high Notch activity and Hes5 expression. *Eur. J. Neurosci.*, **25**, 1006–1022.

38. Giachino, C., Basak, O. and Taylor, V. (2009) Isolation and manipulation of mammalian neural stem cells in vitro. *Methods Mol. Biol.*, **482**, 143–158.