博士論文番号：1081209

The role of CHD7 in human central nervous system development

（ヒト中枢神経系形成過程における **CHD7** の役割）

Chai Muh Chyi

奈良先端科学技術大学院大学

バイオサイエンス研究科　遺伝子発現制御研究室

別所 康全 教授

平成 **30** 年 **1** 月 **10** 日提出

List of contents

List of abbreviations

ATRX	: alpha-thalassemia mental retardation syndrome X-linked
CHD	: chromodomain helicase DNA-binding
CNS	: central nervous system
CRISPR	: clustered regularly interspaced short palindromic repeat
DSB	: double strand break
EGS	: ethylene glycol bis (succinimidyl succinate)
ESC	: embryonic stem cell
FGF	: fibroblast growth factor
HDR	: homology-directed repair
iPSC	: induced pluripotent stem cell
ISW	: The Imitation Switch
NC	: neural crest
NE	: neuroepithelial
NHEJ	: non-homologous end joining
NMD	: nonsense-mediated decay
NPC	: neural progenitor cell
NSC	: neural stem cell
PHD	: plant homeodomain
POU	: the Pituitary-specific Pit-1, the Octamer transcription factor proteins Oct-1 and Oct-2, the neural Unc-86 transcription factor
SANT	: Swi3, Ada2, N-CoR and TFIIIB
shRNA	: short hairpin RNA
SLIDE	: SANT-like ISWI domain
SMA	: smooth muscle actin
SOX	: sry-related HMG box
SWI/SNF	: Switch/Sucrose Non-Fermentable
TSS	: transcription start site

1. Literature Review

1.1. CHARGE syndrome

Back in 1981, CHARGE was first proposed by a group of physicians as the mnemonic for an association of clinical features including C-coloboma, H-heart disease, A-atresia choanae, R-retarded growth and development, G-genital abnormality and E-ear anomalies/deafness, to ease the diagnosis of this disease given its diverse pattern of defects (Pagon et al., 1981). CHARGE syndrome is a congenital disorder with an estimated prevalence of one out of every 10,000 live births. While approximately 70% of the children diagnosed with CHARGE have a life expectancy of up to 5 years old, the highest mortality rate often takes place in the first year of life. Among the most commonly observed dysmorphisms in clinical CHARGE patients are external ear abnormalities (97%); cranial nerve dysfunction (99%); delayed motor development (99%); semicircular canal anomalies (94%); coloboma (81%), congenital heart defects (76%) and severe feeding problems (82%) (Bergman et al., 2011).

The genetic cause of CHARGE syndrome was made known when scientists revealed the presence of mutations in Chromodomain Helicase DNA-binding domain 7 (*CHD7*) gene in a cohort of CHARGE patients (Vissers et al., 2004). The reported *CHD7* mutations are nonsense, splice site or frame-shift, all resulting in either the targeted degradation of aberrant mRNA via nonsense-mediated decay (NMD), or the truncation of CHD7 protein if the mutated transcript escapes NMD. The truncated CHD7 protein is predicted to incur the loss of protein functions, consistent with the clinical observation

that haploinsufficiency of CHD7 accounts for over 90% of CHARGE cases. In accordance with this, homozygous *Chd7* mice are embryonic lethal by embryonic (E) day 11, whereas *Chd7* heterozygous null mice exhibited birth defects that resembles those found in human CHARGE patients, including malformations of the inner ear, eye, choanae, genitals and craniofacial palate.

As for the remaining 5-10% of patients suffered from cardinal clinical features of CHARGE syndrome, but lacking of detectable mutations in *CHD7*, might be succumbed to the limitations of routine genotyping strategies in recognizing multiple aberrant alterations in *CHD7*. In addition, the involvement of other genetic alterations, for instance, mutation in semaphorin 3E (*SEMA3E*) (Lalani et al., 2004) or chromosomal aberrations (Corsten-Janssen et al., 2013; Moustafa-Hawash et al., 2012; Snijders Blok et al., 2014), exposure to teratogen (Komoike et al., 2013), and maternal diabetes are among the suggested alternative causative agents of CHARGE syndrome. However, these alternatives are unlikely to play leading roles compared to *CHD7* due to their sporadic occurrence. Nonetheless, the identification of *CHD7* as the genetic cause of CHARGE syndrome eleven years ago has then immensely facilitated vigorous investigation on the mechanistic actions of CHD7 and its involvement in various developmental pathways, paving the way towards the understanding of pathological mechanism underlying CHARGE syndrome.

1.2. Chromatin Structure and Remodeling

Chromatin is a highly informative complex that contains a large mass of genetic

material compressed into a condensed structure to fit into the nucleus of eukaryotic cells. The packaging of DNA is achieved through nucleosome, the fundamental unit of chromatin (Kornberg and Klug, 1981). Each nucleosome is consisting of 147 base pair (bp) of DNA wrapped in 1.65 turns around a histone octamer. Each histone octamer is comprised of two copies of each histone proteins H2A, H2B, H3 and H4 (Luger et al., 1997). The primary structure of chromatin resembles "beads on a string" model, in which nucleosomes are arranged in a linear array along the DNA. This structure can be further condensed by linker histone, H1, giving rise to higher-order transcriptionally inactive chromatin fibers (Jiang and Pugh, 2009).

Extensive studies revealed that nucleosome remodeling and histone modifications are the principle mechanisms affecting chromatin structures and their underlying roles in genome regulation (Petty and Pillus, 2013). The N terminal tails of histone proteins protruding from the nucleosome are subjected to post-translational modifications, including acetylation, methylation, and phosphorylation. These modifications are likely to alter intra- and inter-nucleosomal interactions and change local chromatin organization (Petty and Pillus, 2013). On the other hand, nucleosome remodeling involves the exchange of canonical histone proteins on nucleosome with histone variants (Kamakaka and Biggins, 2005; Sarma and Reinberg, 2005) and the mobilization of nucleosomes, which define where nucleosomes are positioned with respect to DNA sequences (Struhl and Segal, 2013). Nucleosome positioning are critical to genome regulation, as it physically occlude the access of other DNA-binding proteins to DNA regulatory elements by 10~20 fold compared to nucleosome-depleted regions

(Iyer and Struhl, 1995; Liu et al., 2006). In addition, nucleosome eviction in the gene body downstream of promoter is required for efficient transcriptional elongation (Zhou et al., 2012). Upon the passage of transcribing polymerases, reassembly of nucleosomes is necessary to avoid cryptic transcription (Cairns, 2009).

The ATP-dependent chromatin remodeling complexes are responsible for selective positioning and organization of nucleosomes, by utilizing energy from ATP hydrolysis (Bartholomew, 2014). The five major subfamilies of ATP-dependent chromatin remodelers are SWI/SNF, ISWI, INO80 or SWR1, CHD and ATRX. Interestingly, the different types of remodelers exhibit distinct remodeling mechanisms. CHD subfamily is capable of nucleosome sliding while sensing the length of linker between neighboring nucleosomes that allows them to determine when to stop their action. This property is also referred to as nucleosome spacing (Bartholomew, 2014). Remodelers belong to the SWI/SNF subfamily directly affect the access of TFs to their cognate sites through repositioning, ejection, and unwrapping of nucleosomes (Boeger et al., 2008; Dechassa et al., 2010). Another possibility is observed with INO80/SWR1 complexes, which deposit H2A histone variants into nucleosome flanking the nucleosome-depleted regions (Luk et al., 2010; Mizuguchi et al., 2004; Papamichos-Chronakis et al., 2011). Given their highly coordinated functions and vast influence, the dysregulation of chromatin remodeling has been implicated in cancer, neurological disorders, diabetes and α-thalassemia (Cohen-Carmon and Meshorer, 2012; Dawson and Kouzarides, 2012; Gibbons et al., 2000; Keating and El-Osta, 2012). The following section will focus on CHD proteins, their structures, functions and recently discovered

roles in development, with specific emphasis on CHD7.

1.3. Structure and Function of the CHD Superfamily

In *S. cerevisiae*, CHD1 is the only conserved subunit that exists while there are nine types of CHD proteins (CHD1-CHD9) in human. The additional diversity suggests a tissue-specific expression pattern and context-dependent functions of these proteins in human development (Sims and Wade, 2011). The common denominators of CHD proteins superfamily are a tandem chromodomain and an ATPase domain of the SNF2 superfamily (Hall and Georgel, 2007). The family is demarcated into three subfamilies, each featured by the presence of subfamily-specific sequence motifs. The subfamily I (CHD1 and CHD2), is characterized by an extra DNA binding domain, which preferentially binds to AT-rich sequences (Stokes and Perry, 1995). Besides, it has been shown that human CHD1 recognizes and binds to H3K4me3, a modification indicative of transcriptional activation (Gaspar-Maia et al., 2009). The subfamily II (CHD3/CHD4/CHD5) contains a pair of plant homeodomains (PHDs) in replacement of the chromodomain (Bienz, 2006). CHD3 and CHD4 function as transcriptional repressor by associating with histone deacetylase 1 within the Nucleosome Remodeling and Deacetylase (NuRD) complex, facilitated by the tandem PHD domains (Bowen et al., 2004; Gomez-Del Arco et al., 2016; Xue et al., 1998). The members of subfamily III (CHD6/CHD7/CHD8/CHD9) are homologs of *Drosophila* Kismet protein, which has been shown to regulate body segmentation, Hox gene expression and transcriptional elongation (Sims and Wade, 2011). This subfamily contains an extra SLIDE/SANT

domain and a C-terminal BRK domain, with the function of BRK domain remains unknown. The SANT domain confers binding ability to linker DNA between nucleosomes, allowing a broader association of these remodelers throughout the genome via multiple access points (Petty and Pillus, 2013).

1.4. CHD7 as the Key Regulator in Tissue Development

The *CHD7* gene is located on chromosome 8 and constitutes of 38 exons that spans 188 kb. The expression of *Chd7* was first detected from early embryogenesis into adult stage in a variety of tissues such as brain, ear, olfactory bulb, heart, gut, and craniofacial structures in both mouse and human (Aramaki et al., 2006; Bosman et al., 2005). The dysfunction of Chd7 adversely affect the formation of olfactory bulb, hippocampus and cerebellum in the CNS, the development of inner ear, craniofacial and tracheal and the formation of heart, hinging on the fact that Chd7 is essential to proper mammalian development.

Genome-wide expression profiling in model systems implied that CHD7 could function either as an activator or repressor (Schulz et al., 2014). However, other studies demonstrated that Chd7 occupancy on DNA regulatory elements is correlated with gene activation (Engelen et al., 2011; Schnetz et al., 2009). The latter observation is further corroborated by recent genomic studies revealing that CHD7 binding is largely absent from regions with repressive marks such as H3K9me3 and H3K27me3 (Schnetz et al., 2009), providing evidence against the theory that Chd7 acts as a repressor. The authors also showed that the majority of CHD7 binding sites is located distal to transcription start

sites (TSS), with a subset of them co-localizes with H3K4me1 and P300, and found with open chromatin region that is hypersensitive to DNaseI assay. These findings collectively suggest that CHD7 targets enhancer elements (Schnetz et al., 2009; Schnetz et al., 2010). A drawback of these studies is that the authors failed to provide a functional correlation between CHD7 occupancy and transcriptional regulation.

In the inner ear, conditional deletion of *Chd7* causes cochlear hypoplasia and the complete absence of semicircular canals and cristae. The heterozygous loss of *Chd7* causes the downregulation of FGF signaling and pro-neural gene expression, leading to reduced proliferation of inner ear neuroblasts within the neurogenic domain. The obstructed development of neurogenic domain is also evidenced by the expansion of Tbx1 expression, which is known as the suppressor of neural cell fate (Hurd et al., 2010). This data suggests that Chd7 has critical functions in coordinating the development of inner ear neuroblasts. Recent study demonstrated that young (6 weeks) adult *Chd7*$^{Gt/+}$ mice display compromised functional responses in olfactory sensory neurons when exposed to various odorants. The authors discovered that the olfactory deficit was attributed to a significant reduction in the proliferation of basal cells, the stem cells progenitors in the postnatal olfactory epithelium. As a result, these mice have smaller olfactory bulbs, fewer olfactory sensory neurons and disorganized epithelial ultrastructure. Furthermore, neuronal regeneration following chemical ablation of the olfactory epithelium in *Chd7* deficient mice was either delayed or reduced. Thus, the defects in the proliferation of pro-neuronal basal cells together with the impaired ability in the regeneration of olfactory sensory neurons indicate the importance of Chd7 in regulating neurogenesis during

olfactory bulb development.

In accordance with the aforementioned findings that CHD7 is expressed since early embryonic stage and its expression persists until adult stage, a few studies have shown an enriched expression of CHD7 in actively dividing neural stem cells and progenitors in the subventricular zone (SVZ) and subgranular zone (SGZ) in the hippocampus of adult mice. The inactivation of *Chd7* in adult SGZ neural stem cells results in a loss of stem cell quiescence, a transient increase in cell proliferation and ended up with significantly less neurons in the hippocampus. The study further showed that during the induction of quiescence in NSCs, Chd7 is required to upregulate Hes5 gene expression while negatively regulate cell cycle progression (Jones et al., 2015). Meanwhile, conditional knock out of *Chd7* in adult SVZ neural stem cells show that the depletion of Chd7 leads to a reduction in mature dopaminergic and GABAergic olfactory bulb interneurons mediated by the decreased expression of pro-neural genes Sox4 and Sox11 (Feng et al., 2013). Another member within the Sox family, Sox2, is a well-known regulator in the maintenance of mouse NS/progenitor state and thus neurogenesis (Bylund et al., 2003). Heterozygous loss-of-function mutations in *SOX2* are associated with eye malformations (microphthalmia or anophthalmia) (Williamson et al., 2006). Interestingly, Chd7 was identified as the transcriptional cofactor of Sox2, and co-regulate the expression of *Jag1*, *Gli3* and *Mycn*, the genetic causes of Alagille syndrome, Pallister-Hall syndrome and Feingold syndrome respectively. The authors proposed that the connection between these genetic causes at molecular level is responsible for the overlapping features shared by these syndromes (Engelen et al., 2011).

Using the *Xenopus* system, knockdown of Chd7 during embryogenesis impaired the migration of neural crest, leading to malformed otolith, a part of the vestibular system analogous to the human ear, ocular coloboma, malformations of the craniofacial cartilage; and heart defects, consistent with the major CHARGE clinical features. The effect of Chd7 downregulation on neural crest migration was ascribable to the diminished expression of *Sox9*, *Twist* and *Slug*, the transcription factors crucial to the formation of multipotent, migratory neural crest cells. Moreover, it was found that CHD7 associates with PBAF (polybromo- and BRG1-associated factor-containing complex), forming a chromatin remodeling complex that synergistically activate the neural crest transcriptional circuitry governing the acquisition of multipotency and migratory phenotype (Bajpai et al., 2010). A role for Chd7 in neural crest cell development was also corroborated in *Chd7*$^{Gt/+}$ mice, whereby the conditional deletion of *Chd7* in migrating neural crest cells (*Wnt1-Cre*) perturbed proper craniofacial and tracheal development (Sperry et al., 2014). These discoveries were particularly appealing, given that it was long postulated that CHARGE syndrome is a type of neurocristopathy (Siebert et al., 1985). Taken together, Chd7 has been implicated in several stem cell populations and the functional importance of Chd7 appears to converge on the maintenance of stem cell property and fate specification.

1.5. CRISPR-Cas9 system

The acronym CRISPR (Clustered regularly interspaced short palindromic repeats) was first used to describe the presence of structured loci containing unusual

repeat sequences found across a wide range prokaryotic microbes and archaea (Jansen et al., 2002). Interestingly, scientists found that a portion of these sequences were mapped to viral and bacteriophage genomes, leading to the identification of CRISPR as part of an adaptive immune mechanism against deadly infections (Barrangou et al., 2007; Mojica et al., 2005).

Distinct CRISPR loci are classified into three groups based on the functional type of CRISPR-associated (Cas) protein components employed to locate target sequences and to perform target cleavage. Among these, the type II CRISPR system from *Streptococcus pyogenes*, which utilize only a single effector nuclease, Cas9, emerged as the most commonly used RNA-guided genome editing tools in eukaryotic cells, owing to its simplicity and adaptability (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013a; Mali et al., 2013b). The system requires of two components non-coding RNAs, a crRNA and a trans-activating (tracrRNA). The former guides Cas9 to DNA target sequences via Watson-Crick base pairing while the latter is critical for the maturation of crRNA and recruitment of Cas9 to initiate nuclease activity (Deltcheva et al., 2011). Both RNAs (crRNA and tracrRNA) can be experimentally fused into a single guide RNA (sgRNA), which makes the system even simpler, comprising of just sgRNA and Cas9 (Jinek et al., 2012). The sgRNA and Cas9 components recognize their cleavage targets by a short sequence motif known as the protospacer adjacent motif (PAM), which is located immediately downstream of the desired target site in the genome. In the case of *S. pyogenes*, its activity requires a PAM with sequence 5'-NGG-3', where N can be any nucleotide (Jinek et al., 2012). When Cas9 is recruited to the targeted locus in the genome,

it induces a double strand break (DSB) and triggers host-mediated DNA repair mechanisms, involving either non-homologous end joining (NHEJ) or homology-directed repair (HDR). HDR relies on the availability of repair template to drive high-fidelity repair, a frequently used strategy in vector-based gene knock out and knock-in experiments. Meanwhile, NHEJ at the ends of DSB often causes insertions or deletions (indel) leading to gene disruption (Kim and Kim, 2014).

It is worth to emphasize that CRISPR technology warrants further attention into its versatility. Citing the present study as a case in point, CRISPR was adapted for the addition of epitope tags targeted to endogenous chromatin remodeling factor *CHD7* for chromatin immunoprecipitation followed by next-generation DNA sequencing (ChIP-seq) purposes. Such manipulation has enabled unprecedented characterization of CHD7 genomic binding patterns and thus its functions in human cells, which otherwise will remain poorly defined given the lack of robust and validated antibodies for ChIP-seq. Additionally, the availability of catalytically inactive Cas9 (dCas9) fusion proteins that can either repress (CRISPRi) or activate (CRISPRa) transcription of targeted genes (Gilbert et al., 2013) greatly facilitates genome-wide modulation of transcription of endogenous genes. Such an alternative aspect of CRISPR technology represents a feasible strategy for the study future initiatives to systematically investigate how gene dosage affects cellular functions in multiple congenital disorders and subsequent genetic and drug screenings for factors that mediate such effects.

2. Introduction

Chromatin-based epigenetic alterations are responsible for multiple human neurodevelopmental disorders, including Coffin-Siris syndrome, autism, Alpha thalassemia X-linked intellectual retardation (ATRX) syndrome, Kabuki syndrome, and Rett syndrome (Ronan et al., 2013), suggesting that chromatin regulators play important roles in neural development. CHARGE syndrome is a congenital disorder with multiple features caused by heterozygous mutation of *CHD7* (Vissers et al., 2004). Currently, no effective treatment is available for this disease. CHD7 is a member of the chromodomain helicase DNA-binding (CHD) family, a group of ATP-dependent chromatin remodeling factors that alter chromatin structure by rearranging the position and organization of nucleosomes on DNA (Jiang and Pugh, 2009).

CHARGE syndrome is commonly known as a neural crest (NC)-disease or neurocristopathy in which NC-derived tissues are defective. CHD7-knockdown human embryonic stem cells (ESCs) consistently fail to differentiate into NC-like cells, indicating that CHD7 is required for NC specification and migration (Bajpai et al., 2010). Moreover, several studies of experimental mouse models have reported that Chd7 is essential for neurogenesis (Feng et al., 2013; Jiang et al., 2012; Jones et al., 2015; Layman et al., 2009; Yu et al., 2013) and acts as an upstream activator of the promoters of several pro-neural genes, such as *Sox4, Sox11* and *Gli3* (Engelen et al., 2011; Feng et al., 2013). Moreover, Chd7 plays a pivotal role in the regulation of oligodendrocyte maturation and

myelination (He et al., 2016), substantiating a potentially important function of Chd7 in central nervous system (CNS) development. Given that CHD7 depletion adversely affects the capacity for differentiation towards both neural and NC lineages, it is conceivable that CHD7 is a regulator of cell type-specific gene expression programs. Consistent with this idea, genome-wide chromatin immunoprecipitation followed by sequencing (ChIP-seq) analysis of Chd7 using mouse ESCs revealed that Chd7 regulates the establishment of an ESC-specific gene expression program through binding to enhancer elements, and Chd7 binding preferences change during the transition from ESCs to neural progenitors, indicating that the function of Chd7 varies by developmental stage (Schnetz et al., 2009; Schnetz et al., 2010).

To date, the functional roles of Chd7 have mainly been examined in adult neural stem cells and lineage-committed progenitors from animal models; however, CHD7 is highly enriched in the neural tube, a key structure in neuroectodermal development of the human fetal brain (Sanlaville et al., 2006). Importantly, CHD7 expression is confined to the CNS and mesenchymal structures (Sanlaville et al., 2006), which both originate from the neuroectoderm. Although CNS and craniofacial anomalies frequently co-occur in CHARGE patients (Sanlaville and Verloes, 2007), no study to date has addressed the impact of CHD7 dysfunction on human neuroectodermal development. These deficits in knowledge of the molecular functions of CHD7 and the importance of CHD7-dependent regulation in the etiology of CHARGE syndrome highlight the need for an investigation focused on developmental stages relevant to CHARGE pathogenesis.

In the present study, I used induced pluripotent stem cell-derived

neuroepithelial cells (iPSC-NE cells), which exhibit cellular properties equivalent to those of early NE precursors residing in the neural tube (Falk et al., 2012; Koch et al., 2009), as an *in vitro* model to evaluate the function of CHD7 during neuroectodermal development. By establishing iPSC-NE cells from healthy donors and CHARGE patients, I found that CHD7 plays an essential role in maintaining NE identity and CNS lineage development by indirectly suppressing the induction of the NC. Furthermore, I found that CHD7 controls an epigenetic state that maintains CNS lineage identity, largely through the activation of CNS-specific enhancers. Moreover, I show that CHD7-dependent super-enhancer activation controls the expression of *SOX21* and *BRN2*, which are functionally important in the maintenance of NE identity, as well as for the pathogenesis of CHARGE syndrome.

3. Materials and Methods

3.1. Cell Culture

The human iPSC lines derived from healthy control individuals (201B7, 1210B2 and WD39) and CHARGE syndrome patients (CH1 and CH2) were established and maintained as previously described (Nakagawa et al., 2014; Okita et al., 2013; Okuno et al., 2017; Takahashi et al., 2007). The iPSC-derived neural rosette-like structure and brain organoids were generated from 201B7 as previously described (Isoda et al., 2016a; Lancaster et al., 2013). The AF22 NE cell line was kindly provided by Dr. Austin Smith and maintained as previously described (Falk et al., 2012). The 201B7 and 1210B2 NE cell lines were established from human iPSC lines as previously described (Isoda et al., 2016b). The NE cells were passaged every 3-4 days and plated at a ratio of 1:4. The cells were dissociated using TrypLE^{TM} Select (Thermo Fisher) and plated on Matrigel (Corning)-coated dishes in lt-NES medium (DMEM/F12 (Wako), 2 mM L-glutamine (Nacalai Tesque), 1.6 mg/ml glucose (Nacalai Tesque), 0.1 mg/ml Penicillin/Streptomycin (Nacalai Tesque), N2 supplement (1:100; Thermo Fisher), 1 μl/ml B27 (Thermo Fisher)) or RHB-A medium (Takara) supplemented with 10 ng/ml EGF (Peprotech) and 7.5 μl/ml StemBeads FGF2 (Stem Cultures). The human embryonic kidney (HEK) 293T cell line was grown in high-glucose DMEM (Nacalai Tesque) supplemented with 10% fetal bovine serum (Sigma), 0.1 mg/ml penicillin/streptomycin and 1 mM sodium pyruvate (Sigma).

3.2. Neuronal Differentiation of Human NE cells

The NE cells were first cultured in N2B27 media consisting of a 1:1 ratio of DMEM/F12 and Neurobasal medium (Life Technologies) supplemented with N2 supplement (1:200), GlutaMAX (1:100, GIBCO), 0.8 mg/ml glucose, 0.1 mg/ml penicillin/streptomycin, B27 (1:100) and 3 μM CHIR99021 (Stemgent) for 6 days. At day 6 of CHIR treatment, the cells were passaged and replated onto poly-L-ornithine (Sigma) and laminin (Thermo Fisher)-coated dishes in KBM Neural Stem Cell medium (Kohjin Bio) containing B27 (1:50) supplemented with 1 μM CHIR99021. The next day, the medium was replaced with KBM Neural Stem Cell medium supplemented with 10 ng/ml BDNF (R&D), 10 ng/ml GDNF (R&D), 200 μM ascorbic acid (Sigma) and 500 μM db-cAMP (Sigma) (neurogenic differentiation medium) and cultured for another 9 days. The medium was changed every other day.

3.3. Lentiviral Production and Transduction

To generate CHD7, BRN2 and SOX21 knockdown cells, I transduced human NE cells with lentiviruses expressing shRNA targeting *CHD7*, *BRN2* and *SOX21*, respectively. All shRNA vectors were purchased from Sigma (for details, see Table 1). For rescue experiments, the cDNAs of human *SOX21* [GenBank: NM_007084.2] and *BRN2* [GenBank: NM_005604.3] were cloned into the pENTR-D-TOPO vector (Invitrogen, USA). Using LR reaction (Invitrogen, USA), the cDNAs were inserted into the CSIV-CBh-RfA-IRES2-HygR construct. Lentiviral packaging plasmids pCAG-HIVgp and pCMV-VSV-G-RSV-Rev were co-transfected into HEK293T cells with Gene

Juice (Merck) for virus production. Lentiviral supernatant was collected 48 hours after transfection and supplemented with 4 μg/ml polybrene, filtered through a 0.45-μm filter, and used to infect human NE cells. At 48 hours after infection, the cells were selected with either 1 μg/ml puromycin (Nacalai Tesque) or a combination of 1 μg/ml puromycin and 80 μg/ml hygromycin (Sigma).

Table 1. List of constructs used in this study

Vectors	**Source**	**Identifier**
pLKO.1-puro Non-Target shCTRL	Sigma	SHC016
pLKO.1-puro shCHD7 (KD1)	Sigma	TRCN0000016410
pLKO.1-puro shCHD7 (KD2)	Sigma	TRCN0000016411
pLKO.1-puro shBRN2 (KD1)	Sigma	TRCN0000019330
pLKO.1-puro shBRN2 (KD2)	Sigma	TRCN0000230048
pLKO.1-puro shSOX21	Sigma	TRCN0000420797
CSIV-CBh-RfA-IRES2-HygR	Miyoshi et al. 1998	N/A
pCAG-HIVgp	Gift from Dr. Hiroyuki Miroshi	N/A
pCMV-VSV-G-RSV-Rev	Gift from Dr. Hiroyuki Miroshi	N/A
pENTR-D-TOPO	Thermo Fisher Scientific	240020
pcDNA-Flag-His-CHD7	Gift from Dr. Joanna Wysocka	N/A
pCMV-hyPBase	Yusa et al. 2011	N/A
PiggyBac Dual Promoter Vector	System Biosciences	PB513B-1
pENTR-CHD7-wt	This study	N/A
PB513-CHD7-wt	This study	N/A
PB513-CHD7G1391fs	This study	N/A
PB513-CHD7R1494X	This study	N/A
pX335	Addgene	42335
pGL3-Promoter Vector	Promega	E176A
pEF-Renilla-Luc	Nakashima et al. 1999	N/A

3.4. Construction of plasmids and the generation of stably transfected NE cell lines

pcDNA harboring the full-length human *CHD7* transcript (Flag-His-*CHD7* fusion transcript) was a kind gift from Dr. Joanna Wysocka. The *CHD7* transcript was first cloned into the pENTR-D-TOPO vector (pENTR-*CHD7*-wt) before inserted into the PiggyBac vector PB513B-1 (System Biosciences) using LR reaction (PB513-*CHD7*-wt). Mutant *CHD7* (*CHD7*^{G1391fs} and *CHD7*^{R1494X}) were generated using a modified PCR-based site-directed mutagenesis strategy (Higuchi et al., 1988) using full-length wild-type *CHD7* as a template. Two primer sets (primers are listed in Table 2) carrying the desired point mutations and one outer primer set were used to PCR amplify fragments of *CHD7* from the pENTR-*CHD7*-wt vector. PCR products were cloned into the PB513-*CHD7*-wt vector to replace the corresponding regions of wild-type *CHD7*.

Stable transfection of AF22 cells and CHARGE-NE cells with either mutant CHD7 proteins or wild-type CHD7 was performed using a Nucleofector device (Amaxa) following the manufacturer's guidelines. The PB513-*CHD7*^{G1391fs}, PB513-*CHD7*^{R1494X} and PB513-*CHD7*-wt vectors were co-electroporated with pCMV-hyPBase vector (Yusa et al., 2011) at a ratio of 2.5:1 into the respective NE cell lines. Control cell lines were generated using an empty PB513B-1 vector. Stable clones were selected with puromycin (1 μg/mL) 48 hours after electroporation. Resistant cells were selected for at least 3 passages before further analysis.

Table 2. List of primers used in this study

sgRNA construction	**Primer sequences (5' ---> 3')**
pX335_CHD7_HA_F	CACCGACTTGAACTGGAACTGGTAC
pX335_CHD7_HA_R	AAACGTACCAGTTCCAGTTCAAGTC
SURVEYOR assay	**Primer sequences (5' ---> 3')**
Sur_CHD7_HA_Fw2	CCTGGATAAGACTGCAGAGTCC
Sur_CHD7_HA_Rv2	TGCACTGCACAATACTTAATGACC
qPCR	**Primer sequences (5' ---> 3')**
CHD7-N_Fw	TTCCATCACCACCCCTCTAC
CHD7-N_Rv	GACTGTCTGGCTCCGAGAAC
CHD7-C_Fw	CTTTTCATGAGCCACAAACG
CHD7-C_Rv	TCTTCTTCAAAAGCTTTGGTCAC
SOX21_Fw	GTGGAAACTGCTCACAGAGTC
SOX21_Rv	CACCGGGAAGGCGAACTTGTC
BRN2_Fw	CAGCGCACCTCCCACCCTCCATG
BRN2_Rv	GTCGGCATGGTGTGGCTCGTC
SOX9_Fw	CCGGCATGAGCGAGGTGCACTC
SOX9_Rv	CTCTCGCTTCAGGTCAGGTCAGCCTTGC
TWIST1_Fw	CATCCCCACGCTGCCCTCGGAC
TWIST1_Rv	CTCCATCCTCCAGACCGAGAAG
SNAI2_Fw	CCTCCAAAAAGCCAAACTACAGCGAAC
SNAI2_Rv	CTGAGGATCTCTGGTTGTGGTATGAC
MSX1_Fw	GCTCGGCCATTTCTCGGTGGGGGGACT
MSX1_Rv	GTCTTGTGTTTGCGGAGGGTGCAGGC
MSX2_Fw	AATTCAGAAGATGGAGCGGCG
MSX2_Rv	GTGCAGGTGGTAGGGCTCATATGTC
Luciferase reporter assay	**Primer sequences (5' ---> 3')**
SOX21 R1_Fw (NheI)	atagctagcGTGTGTGGGAAGCAGCAGAA
SOX21 R1_Rv (XhoI)	atactcgagTGTAGTGCCAACATACAATGCAA
SOX21 R2_Fw (NheI)	atagctagcGAAGATGGTACTTACACTTGGCAGTAGG
SOX21 R2_Rv (XhoI)	atactcgagCTTGCTCTCTGTGTGATGCTGGAGGTC

SOX21_R3_Fw (KpnI)	ataggtaccCTAGGTATCCCCCAGCCCCTACCTTTAT
SOX21 R3_Rv (SacI)	atagagctcGCGCTGTGTAAGCAGTAAGAACTGGTGCTG
SOX21 R4_Fw (KpnI)	ataggtaccCTCCTTCCTTAGTTCCTCCATCAG
SOX21 R4_Rv (SacI)	atagagctcGAGGCTTGCTAGGTGCTGAGATAC
BRN2 R1_Fw (NheI)	atagctagcCCAACTTGCCCCATTTCTATCT
BRN2 R1_Rv (XhoI)	atactcgagCCAACCATTCAAAAATGACATGA
BRN2 R2_Fw (NheI)	atagctagcCTAATTCCTTGACATCTACAGAC
BRN2 R2_Rv (XhoI)	atactcgagGAAGTTTTATGGGCAATAAGGCTAC
BRN2 R3_Fw (KpnI)	ataggtaccGTCCCTACAAATAAAATCTTAAC
BRN2 R3_Rv (NheI)	atagctagcGAGAGCAGTTAACCACAGTCG
BRN2 R4_Fw (SacI)	atagagctcCTAAAGTCAGTTCCCTAGAGAAG
BRN2 R4_Rv (XhoI)	atactcgagGAACGTAGTGTCTTTTCTGGC
Mutant CHD7 construction	**Primer sequences (5' ---> 3')**
CHD7_G1391fx_Fw	ATTCAGACTGGAATCCCAAAATGACCTCCAGGCT
CHD7_G1391fx_Rv	AGCCTGGAGGTCATTTTGGGATTCCAGTCTGAAT
CHD7_R1494X_Fw	ATCAGATCCTCCTACGTTGAACCCACACCATTACC
CHD7_R1494X_Rv	GGTAATGGTGTGGGTTCAACGTAGGAGGATCTGAT
CHD7_outer_Fw2	CCTTATCAATGGTGCTGAAGAG
CHD7_outer_Rv2	CAGGGTCATTGAGGATGTGATAATC

3.5. Immunocytochemistry

Cells were washed with PBS, fixed in 4% paraformaldehyde at room temperature for 15 min and rinsed three times with PBS. After the cells were blocked with PBS solution containing 3% FBS and 0.1% Triton X-100, they were incubated with primary antibodies at 4°C overnight. Fluorescent marker (Cy2, Cy3, or Cy5)-conjugated secondary antibodies (Jackson Laboratory) were used to label each primary antibody at room temperature for 2 hours. Nuclei were stained with H33258 (Nacalai Tesque). Samples were viewed and photographed using either a Zeiss Axioplan 2 fluorescence

microscope or Zeiss Confocal Laser Scanning Microscope (LSM 700). Detailed information for primary antibodies is summarized in Table 3.

Table 3. List of primers used in this study

Antibodies used for Immunostaining	Source	Identifier
Goat polyclonal anti-SOX1	R&D	AF3369
Mouse monoclonal anti-SOX2	Millipore	MAB2018
Rabbit polyclonal anti-CHD7	Bethyl	A301-223A
Rabbit monoclonal anti-CHD7	Cell Signaling	D3F5
Mouse monoclonal anti-Nestin	Millipore	MAB5326
Chick polyclonal anti-NeuN	Aves	NUN
Rabbit polyclonal anti-βIII-tub	Bio Legend	802001
Rabbit polyclonal anti-DACH1	Proteintech	10914-1-AP
Rabbit polyclonal anti-AP2a	Cell Signaling	3215S
Mouse monoclonal anti-AP2a	Santa Cruz	sc-12726
Rabbit polyclonal anti-p75	Promega	G3231
Rabbit polyclonal anti-Zic1/2	Millipore	ABE1958
Mouse monoconal anti-SMA	Sigma	A2547
Rat monoclonal anti-HA	Roche	3F10
NL557-conjugated Goat anti-human SNAIL	R&D	SC026
Rabbit polyclonal anti-SOX9	Santa Cruz	sc-20095
Chick polyclonal anti-GFP	Aves	GFP-1020
Mouse monoclonal OCT3/4	Santa Cruz	sc-5279
Antibodies used for ChIP		
Rabbit polyclonal anti-CHD7	Bethyl	A301-223A
Rabbit monoclonal anti-CHD7	Cell Signaling	D3F5
Rat monoclonal anti-HA	Roche	3F10
Mouse monoclonal anti-H3K27ac	MBL	MABI0309
Mouse monoclonal anti-H3K4me1	MBL	MABI0302

Mouse monoclonal anti-H3K4me3	MBL	MABI0304
Mouse monoclonal anti-H3K27me3	MBL	MABI0323
Mouse monoclonal anti-H3K9me3	MBL	MABI0308
Mouse monoclonal anti-H3K36me3	MBL	MABI0331
Dynabeads M-280 sheep anti-mouse IgG	Thermo Scientific	11202D
Dynabeads M-280 sheep anti-rabbit IgG	Thermo Scientific	11204D
Dynabeads M-280 sheep anti-rat IgG	Thermo Scientific	11035
Antibodies used for Simple western		
Rabbit monoclonal anti-CHD7 (N-terminus)	Cell Signaling	D3F5
Rabbit polyclonal anti-CHD7 (C-terminus)	Abcam	ab31824
Rat monoclonal anti-HA	Roche	3F10
Rabbit monoclonal LaminB1	Cell Signaling	12586

3.6. Immunohistochemistry

E10.5 whole embryos were washed with PBS, fixed in 4% paraformaldehyde at room temperature for 30 min and rinsed three times with PBS. Whole embryos were cryoprotected in 30% sucrose in PBS overnight at 4°C and then embedded in OCT compound. Cryostat sections (12 μm) were cut and affixed to APS-coated glass slides (Matsunami Glass). The sections were then blocked with PBS containing 5% FBS, 1% albumin and 0.25% Triton X-100 for 1 hour at room temperature. The sections were incubated with primary antibodies at 4°C overnight. Fluorescent marker (Cy2, Cy3, or Cy5)-conjugated secondary antibodies (Jackson Laboratory) were used to label each primary antibody at room temperature for 2 hours. Optical sections were viewed using a

Zeiss Confocal Laser Scanning Microscope (LSM 700) with 20x and 40x objectives. Detailed information for primary antibodies is summarized in Table 3.

3.7. CRISPR-Cas9-mediated protein tagging

In general, the CRISPR-Cas9 system was exploited as previously described (Cong et al., 2013; Ran et al., 2013). Plasmids expressing wild-type Cas9 under the control of the CBh promoter and a human U6 promoter-driven sgRNA cloning vector were purchased from Addgene (http://www.addgene.org/). sgRNA was designed to target sequences immediately downstream of the stop codon of *CHD7* using the CRISPR tool (http://crispr.mit.edu). sgRNA was constructed using the primers listed in Table 2. The SURVEYOR assay was performed to examine the targeting efficiency and specificity of gRNA. The cleavage efficiency was examined by transfecting the plasmid in HEK293T cells, followed by the SURVEYOR assay. Briefly, genomic DNA was extracted using a DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions, and CRISPR-targeted sites in genomic regions were amplified with PrimeSTAR Max DNA polymerase (Takara). PCR products were subjected to the SURVEYOR assay following the manufacturer's protocol (Transgenomics). The primers used are listed in Table 2. To facilitate HDR, I purchased a single-stranded DNA oligo donor (ssODN) consisting of a homology sequence, HA tag sequence, and an EcoRI restriction site for screening purposes from Integrated DNA Technologies. The ssODN was designed to include the removal of a single nucleotide within the protospacer adjacent motif (PAM) to eliminate multiple genome editing events at the same loci.

To generate HA knock-in NE cells, sgRNA and Cas9 expression plasmids were electroporated together with a ssODN at a ratio of 1:1 into AF22 cells using a Nucleofector device (Amaxa) following the manufacturer's guidelines. The electroporated cells were cultured and maintained before FACS sorting. The Single-cell sorting was performed using a flow cytometer SH800 (Sony), and the cells were subsequently cultured in RHB-A medium (Takara) supplemented with EGF (Peprotech) and FGF2 beads (Stem Cultures) in a Matrigel (Corning)-coated 96-well plate (Greiner Bio-One). Each clone derived from a single cell was expanded for further analysis.

3.8. Simple Western

Cells were scraped and collected from 10 cm culture dishes in 300 μl of ice-cold NP-40 lysis buffer with 1X protease inhibitor cocktail (Roche). After the cells were incubated at 4°C for 10 min, they were homogenized by being drawn through a 23-gauge needle to rupture the plasma membrane. The soluble cytoplasmic fraction was separated from the nuclear fraction by centrifugation at 800 ×*g* for 10 min at 4°C. The remaining pellet was washed once with ice-cold lysis buffer without NP-40, followed by centrifugation. Subsequently, the pellet was resuspended in ice-cold RIPA lysis buffer with 1X protease inhibitor cocktail. The lysate was sonicated three times with 30-s pulses (30-s interval between pulses) and cleared of cell debris by centrifugation at 20,000 ×*g* for 10 min. The supernatant was collected as a nuclear fraction. The protein concentration was estimated using the BCA method with a BCA protein assay kit (Pierce) and iMark microplate absorbance reader (Bio-Rad). I used Simple Western system (Simon™,

ProteinSimple), a non-gel-based Western blot-like substitute, to analyze protein expression. The preparation of samples, antibodies and loading of reagents was performed as instructed by the ProteinSimple user manual. Automated separation electrophoresis and chemiluminescence detection were performed using a ProteinSimple WES instrument. The resulting digital images were analyzed with Compass software (ProteinSimple). Detailed information for primary antibodies is summarized in Table 3.

3.9. RNA Isolation and Quantitative RT-PCR

Total RNAs were extracted using a RNeasy Mini Kit (QIAGEN). cDNAs were synthesized from 1 μg total RNA with a ReverTra Ace qPCR RT Kit (TOYOBO) according to the manufacturer's instructions. qPCR analysis was performed using a ViiA7 system (Applied Biosystems) with Fast SYBR Green Master Mix (Thermo Fisher). The expression of target genes was normalized to that of glyceraldehyde 3-phosphate dehydrogenase. The primers used are listed in Table 2.

3.10. Microarray and data analysis

RNA quality was assessed with an Agilent RNA 6000 Nano kit and processed on an Agilent 2100 Bioanalyzer. All RNA samples were amplified and labeled using a TargetAmp-Nano Labeling Kit (Epicentre). The samples were hybridized to Illumina HumanHT-12 v4 Expression BeadChip arrays. Biological duplicates were collected for each sample. Raw probe intensity data were exported from Illumina GenomeStudio gene expression software (v 1.9.0) and loaded onto R (v3.2.2) statistical computing software

for statistical analysis. Limma software (v 3.26.20) was used for background correction, quantile normalization and log (base 2) conversion. Gene sets with no detectable expression in all samples were excluded after normalization. For the evaluation of the overall gene expression profiles, genes differentially expressed by 1.5-fold between in control versus CHD7-knockdown samples and healthy control versus CHARGE-NE samples were extracted and used for GO analysis using DAVID Bioinformatics Resources (http://david.ncifcrf.gov).

Correlations of differentially expressed genes following CHD7 knockdown versus iPSC-derived NPCs and NCCs were evaluated using ExAtlas (https://lgsun.irp.nia.nih.gov/exatlas/). The criteria of significance for the correlation analysis were adjusted with FDR < 0.05 and change > 2-fold. To plot the correlation in a heat map, datasets were sorted with hierarchical clustering. Published gene expression profiles of iPSC-derived NPCs and iPSC-derived NCCs were loaded from the GEO database. The datasets used in study for comparison were as follows: iPSC-derived NPCs (GSM1553290, GSM1553289, GSM1553291, GSM1538561, GSM1538558, GSM1538560) and iPSC-derived NCCs (GSM1470884, GSM1470883, GSM1470885, GSM1538548, GSM1538547, GSM1538546).

3.11. ChIP Assay and Preparation of ChIP-Seq libraries

Cells grown on 10 cm dishes (~9, 000,000 cells; 3 × 10 cm dishes) were used for each ChIP reaction. For single fixation with formaldehyde, cells were fixed with 1% formaldehyde-containing medium for 10 min. For double fixation, the cells were first fixed with 2 mM ethylene glycol bis (succinimidyl succinate) (EGS) for 30 min followed by with 1% formaldehyde-containing medium for 10 min. All fixation steps were performed at room temperature. Cells were washed and incubated with NP-40 buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, and 0.5% NP-40) for 10 min at RT with agitation. The buffer was aspirated and cells were scraped, collected in fresh NP-40 buffer and centrifuged at 1000 ×*g* for 5 min at 4°C. The pellet was resuspended in SDS lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS and protease inhibitor cocktail) and topped up with 400 μl of ChIP dilution buffer (50 mM Tris-HCl pH 8.0, 167 mM NaCl, 1.1% Triton X-100, 0.11% sodium deoxycholate and protease inhibitor cocktail). The lysates were sonicated with either a Bioruptor (COSMO BIO) (treatment time: 15 min; 30 sec ON, 60 sec OFF, 10 cycles, power high) or a Covaris S2 focused-ultrasonicator (intensity: 4; duty cycle: 5%; cycles per burst: 200; treatment time: 15 min) followed by centrifugation at 20,000 ×*g* at 4°C for 10 min.

For each ChIP reaction, 50 μl of Dynabeads M-280 sheep anti-mouse IgG, Dynabeads M-280 sheep anti-rabbit IgG, or Dynabeads sheep anti-rat IgG (all from Life Technologies) were washed with PBS twice. Beads were collected by magnet and conjugated to 5 μg of desired primary antibody in 500 μl of ice-cold RIPA buffer I (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100 and

0.1% sodium deoxycholate) with 1x protease inhibitor cocktail. The suspensions were rotated at 4°C overnight. The next day, beads were collected by magnet and washed twice with 500 μl of ice-cold RIPA buffer I. Following the last wash, sheared chromatin was added to the magnetic beads and allowed to react overnight at 4°C with rotation. The immunoprecipitated samples were washed once with RIPA buffer I, once with RIPA buffer II (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100 and 0.1% sodium deoxycholate), and twice with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Each wash was accomplished with rotation at 4°C for 5 min. Immunoprecipitated DNAs or input DNAs were eluted by adding 200 μl of ChIP elution buffer (10 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM EDTA, 0.5% SDS) and were incubated at 65°C overnight to reverse the crosslinks. Samples were then treated with RNase A at 37°C for 30 min and with Proteinase K at 55°C for 3 h. DNA was purified using a ChIP DNA Clean & Concentrator Kit (Zymo Research). ChIP-seq libraries were prepared from 3-5 ng of ChIP-enriched DNA using KAPA Hyper Prep Kit (Kapa Biosystems) and sequenced using a HiSeq 2500 (Illumina). Antibodies used for ChIP-seq experiments are listed in Table 3.

3.12. ChIP-Seq Peak-Calling and Data Analysis

Sequenced raw reads were trimmed based on read length and read quality using Trimmomatic (v0.33). The trimmed reads were aligned to the reference genome (UCSC hg19) using Bowtie2 (v2.1.0) with the default parameters, and only uniquely mapped reads were used for downstream analyses. The resulting SMA files were converted to the

BAM format using SAMtools (v0.1.19). Peak calling was performed using Model-based Analysis of ChIP-seq (MACS2) version 2.1.0.20150731 with default setting (Zhang et al., 2008). Overlaps for each peak were calculated using bedtools (v2.17.0) with default parameters. The genome-wide peak distribution from TSS was calculated using the ChIPseeker R package, and functional analysis (GO) of CHD7 peaks was performed using the GREAT website (http://bejerano.stanford.edu/great/public/html/index.php). Heatmaps for each sample against the CHD7 binding region were made using deepTools (v1.5.11). The H3K27ac density plot at CHD7 binding regions, CNS-specific enhancers, TEs and SEs were created using deepTools. The average H3K27ac read density at each region and the corresponding flanking region were calculated (bin size=50). The length of typical and SE regions (between Start and End) was scaled relative to its median length. The enhancer category was calculated based on overlap of the CHD7 binding site and the enhancer region; these overlaps were validated in the VISTA enhancer database.

3.13. Identification of super-enhancers and associated genes

SEs were identified using the ROSE algorithm with default parameters based on H3K27ac intensity in which enhancer peaks located within 12.5 kb were stitched together and ranked based on their input-subtracted signal of H3K27ac (Loven et al., 2013; Whyte et al., 2013). I used a promoter exclusion zone of 5,000 bp to exclude any enhancer contained within a window of ± 2,500 bp around an annotated transcription start site from being stitched. Enhancer-associated genes were defined on the basis of the calculated distance from the center of the super-enhancer to the nearest TSS of each gene.

3.14. Reporter cloning and luciferase assay

Selected genomic regions spanning the human *SOX21* and *BRN2* genes were amplified from the genomic DNA of AF22 cells and cloned into a pGL3-promoter vector (Promega). Primers used for cloning and the coordinates corresponding to CHD7 peaks are listed in Table 2. Reporter vectors were co-transfected with pEF-Renilla-luc (Nakashima et al., 1999) at a ratio of 20:1 using Viafect transfection reagent (Promega) into AF22 cells transduced with either control or *CHD7*-shRNA pLKO-lentiviral vectors. Media was replaced with fresh media after overnight incubation, and luminescence was measured 48 hours after transfection using a Dual-Luciferase Reporter Assay System (Promega). Firefly luciferase activity was normalized to Renilla luciferase activity and expressed as fold-change relative to control knockdown samples.

3.15. Statistical analysis

All analyses presented in this study were conducted using GraphPad Prism 7 (San Diego, California). Values are expressed as the mean ± SEM. When comparing two groups, statistical significance was determined using unpaired, two-tailed Student's *t* test. When comparing multiple groups, one-way ANOVA was performed with multiple comparisons, followed by Bonferroni analysis.

4. Results

4.1. NE cells express higher levels of CHD7 than NC cells

To explore the function of CHD7 in NE cells, I initially examined the expression of CHD7 in iPSC-derived neural rosettes (Curchoe et al., 2012), which are highly organized multicellular structures of NE cells. As shown in Fig. 1, I observed CHD7 expression in SOX1- and SOX2-positive neural rosettes. I also observed intense CHD7 expression in most SOX1- and Nestin-positive NE cells (Fig. 1). Following neuronal differentiation of NE cells, CHD7 expression remained enriched in βIII-tubulin-positive neurons but was faintly detectable only in NeuN-positive mature neurons (Fig. 1), consistent with previous *in vivo* data showing that the expression of *Chd7* is turned off in mouse hippocampal and cerebellar granule cells (Feng et al., 2017; Habib et al., 2016; Jones et al., 2015). I further examined the expression of CHD7 in brain organoids derived from iPSCs (Lancaster et al., 2013) and observed that CHD7 expression was decreased in NeuN-positive neurons (Fig. 1). These results suggest that the expression of CHD7 is functionally required before terminal differentiation of NE cells. Given the morphological and structural resemblance between the neural rosette and embryonic neural tube, CHD7 expression in NE cells recapitulates the *in vivo* expression of CHD7 in the neural tube of human fetal brains (Sanlaville et al., 2006).

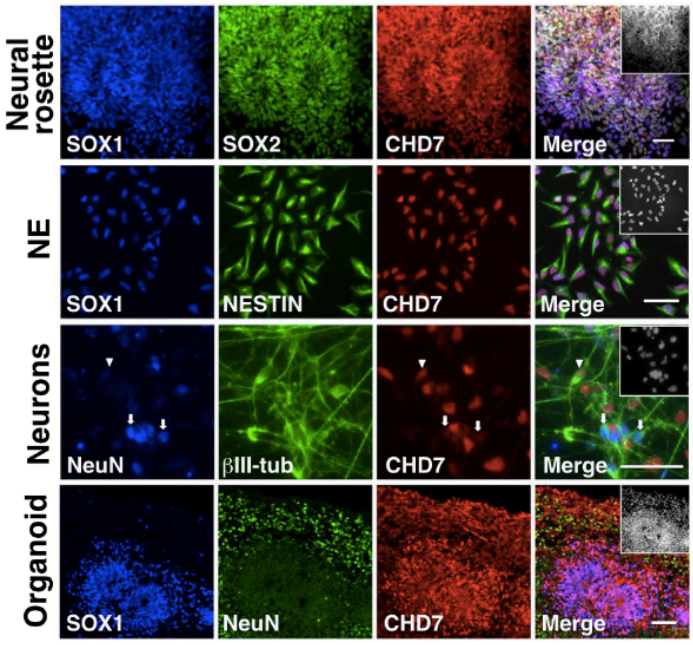


Figure 1: CHD7 is highly expressed in NE cells.
Immunocytochemistry showing that human iPSC-derived neural rosettes at day 5 of neural induction (first row), and that the neural rosettes expanded as monolayer cultures (second row), NE-derived neuronal cells (third row) and 60-day-old human iPSC-derived brain organoids (fourth row). Cells were stained with antibodies against SOX1, SOX2, NESTIN, CHD7, NeuN and βIII-tublin. Arrows indicates mature neuron while arrowhead indicates immature neurons. Insets: Hoechst nuclear staining of each field. Scale bars, 50 μm.

Since CHARGE syndrome is commonly considered a neurocristopathy, and CHD7 is required for the formation of the migratory NC (Bajpai et al., 2010), I next sought to compare the expression levels of CHD7 between iPSC-derived AP-2α-positive NC cells and NE cells. The CHD7 expression level was lower in NC cells than in NE cells (Fig. 2). I further sought to compare the expression level of Chd7 between NC and NE cells by performing immunohistochemistry in mouse embryonic day (E) 10.5 neural tube sections. *In vivo*, the expression level of Chd7 in Sox1-positive NE cells was visibly higher than that in p75- or Zic1/2-positive migrating NC cells (Achilleos and Trainor, 2012; Simoes-Costa and Bronner, 2013) located outside the neural tube (Fig. 3).

Taken together our *in vivo* and *in vitro* data, I hypothesized that higher level of CHD7 in NE cells define NE fate.

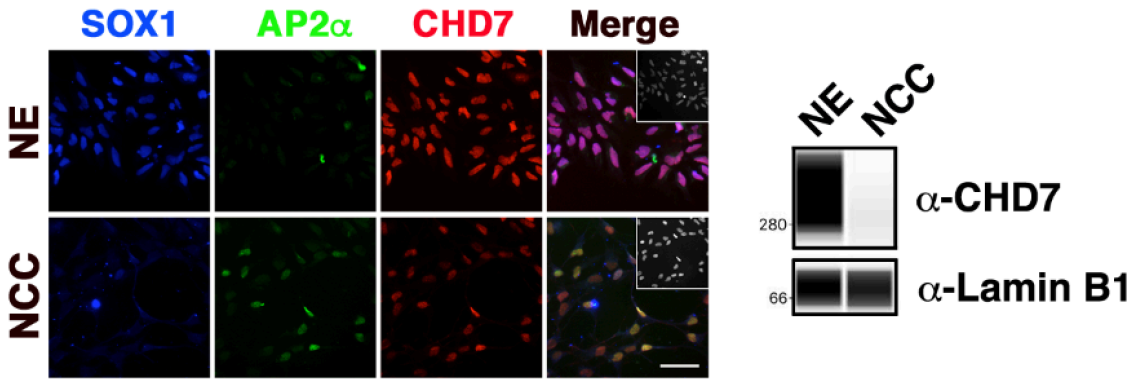


Figure 2: CHD7 is highly expressed in NE cells than NC cells.
Immunocytochemistry of NE cells and human iPSC-derived neural crest cells (NCCs) for SOX1 (blue), AP-2α (green) and CHD7 (red). Insets: Hoechst nuclear staining of each field. Scale bars, 50 μm. Immunoblot analysis of CHD7 expression in NE cells and NCCs. Lamin B1 was used as a nuclear loading control (right panel).

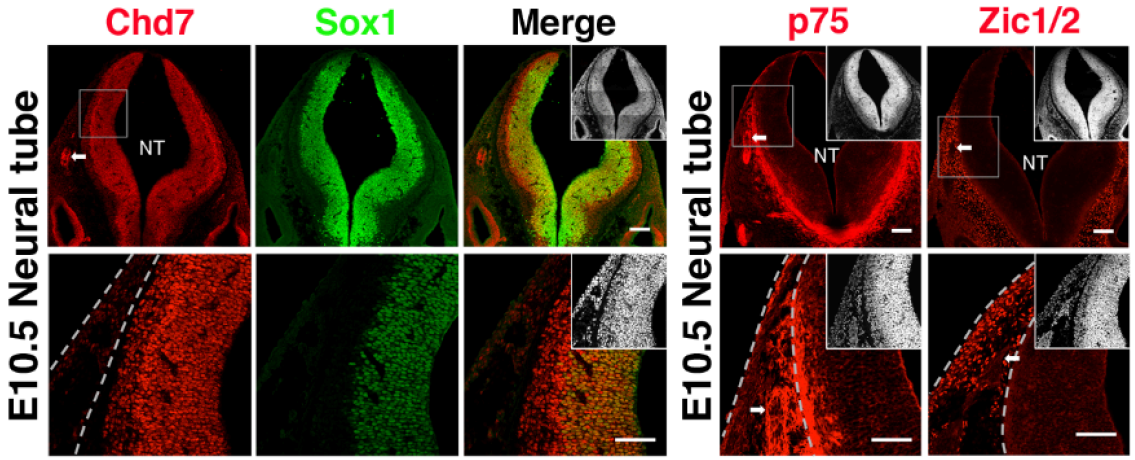


Figure 3: Chd7 expression was detected specifically in E10.5 mouse neural tube.
Immunostaining for Chd7 and Sox1 of E10.5 neural tube sections (upper-left panel). Immunostaining for Chd7 and migratory neural crest markers (p75 and Zic1/2) of E10.5 neural tube sections (upper-right panel). The small box in the upper panels denotes a magnified region shown in the lower panels. The arrow points to a neurogenic dorsal root ganglion (DRG) clusters. Dotted lines indicate the dorsolateral and ventrolateral neural crest cell migratory pathways. Insets: Hoechst nuclear staining of each field. Scale bars, 100 μm. NT=neural tube.

4.2. CHD7 is required for the maintenance of neuroepithelial identity

Next, to ascertain whether the expression level of CHD7 defines NE fate, I conducted lentiviral short hairpin RNA (shRNA)-mediated knockdown of CHD7 in NE cells (AF22 cell line). Two different shRNAs against *CHD7* were used, and the knockdown efficiency of CHD7 was validated by immunocytochemistry and quantitative RT-PCR (qRT-PCR) (Fig. 4).

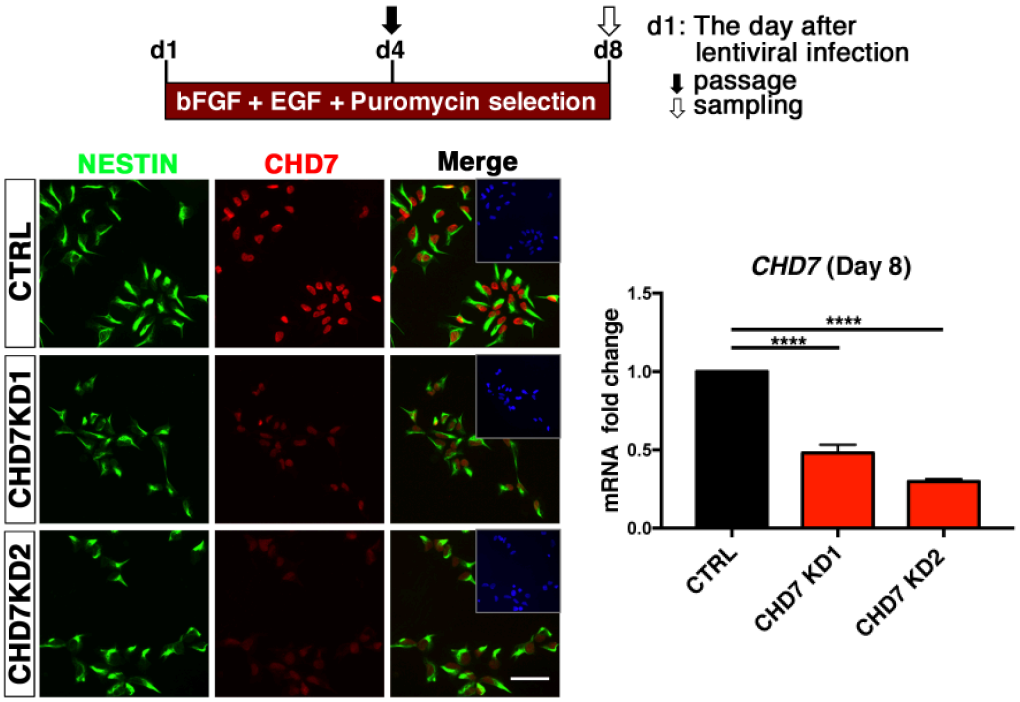


Figure 4: Lentiviral-mediated knockdown of CHD7.
Schematic diagram illustrating fixation and analysis of CHD7-knockdown cells (upper panel). Immunostaining for Nestin (green) and CHD7 (red) in control and CHD7-knockdown NE cells at day 8 after lentiviral transduction (lower panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. qRT-PCR validation of shRNA-mediated downregulation of *CHD7* mRNA at day 8 (right panel) (n=3). Data are presented as the mean ± SEM (****$p < 0.0001$, two-tailed unpaired Student's t test).

In the presence of *CHD7* shRNAs, the number of cells expressing definitive NE markers, including SOX1 and DACH1 (Koch et al., 2009), was markedly reduced, indicating that knockdown of CHD7 resulted in the loss of NE identity (Fig. 5).

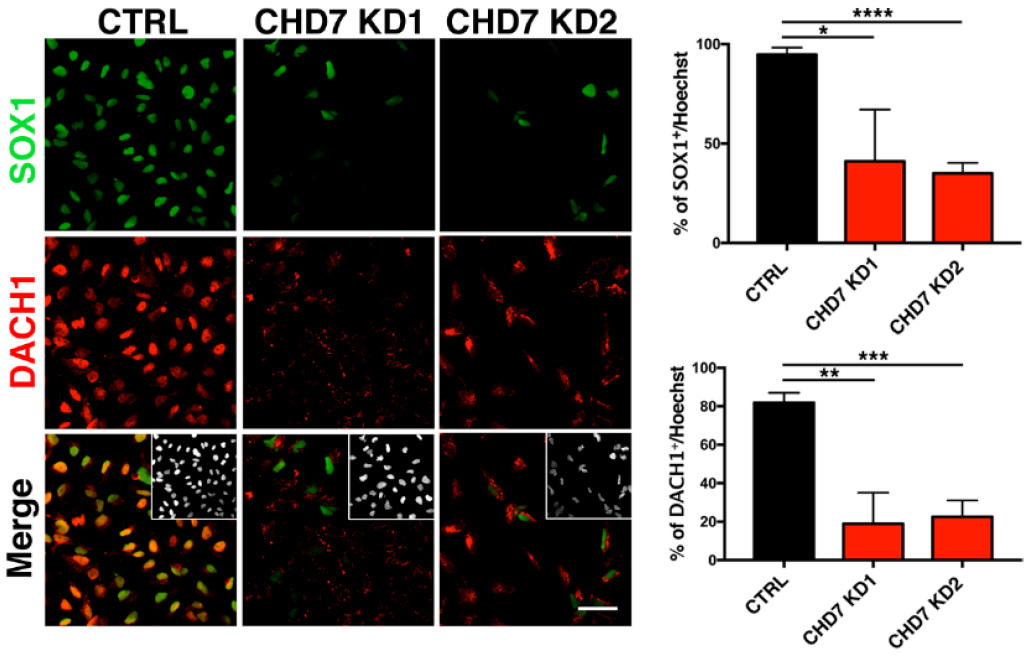


Figure 5: The loss of definitive NE markers expression following CHD7 knockdown.
Immunocytochemistry for SOX1 and DACH1 expression in NE cells (lower panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 μm. Quantification of the number of SOX1- and DACH1-expressing cells is shown in the right panel (n=3). Data are presented as the mean ± SEM (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, two-tailed unpaired Student's t test).

Since previous studies have shown that NE cells are highly neurogenic and predominantly give rise to neurons (Koch et al., 2009), I then addressed whether CHD7-knockdown cells retain their neurogenic potential despite the loss of NE identity. To accelerate the neuronal differentiation of transduced NE cells, I treated the cells with CHIR-99021 (CHIR), a small-molecule inhibitor of glycogen synthase kinase 3 (GSK-3), prior to induction of terminal neuronal differentiation (Fig. 6) (Li et al., 2011; Shimojo et al., 2015). NE cells carrying control shRNA differentiated into βIII-tubulin-positive neurons, and some co-expressed NeuN (Fig. 6). By contrast, both βIII-tubulin- and NeuN-positive neurons derived from CHD7-knockdown NE cells were decreased (Fig. 6). Interestingly, concomitant with the loss of neurogenic potential, CHD7 knockdown induced the emergence of smooth muscle actin (SMA)-expressing cells under neurogenic

conditions (Fig. 6); this identity is at odds with NE cells, which are classically known as CNS stem cells that mainly differentiate into neurons, astrocytes and oligodendrocytes. When stimulated, NE precursor cells are capable of redirecting their commitment to the NC lineage in a temporally dependent manner in avian embryos (Scherson et al., 1993). Thus, I speculated that the emergence of SMA-expressing cells instead of neuronal differentiation after CHD7 knockdown was a result of a fate switch to the NC lineage.

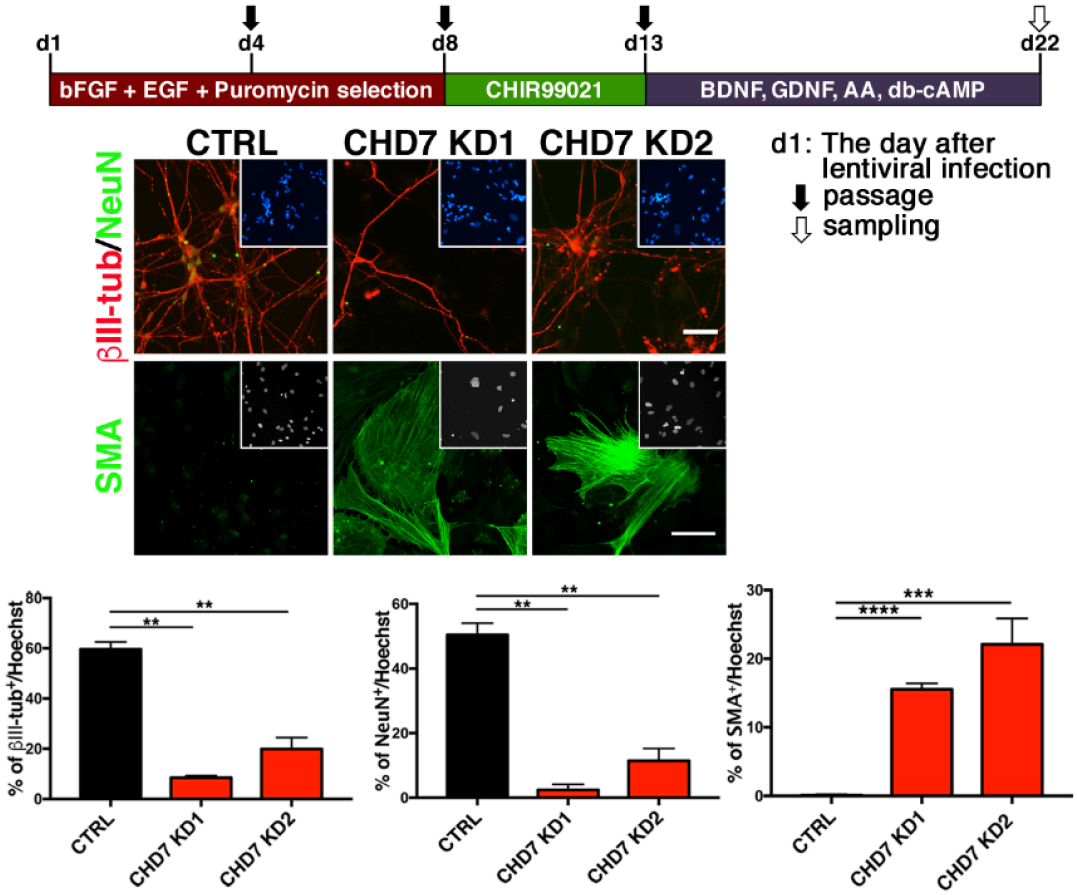


Figure 6: Altered neuronal differentiation ability of CHD7-knockdown cells.
Schematic of the neuronal differentiation protocol (upper panel). Immunostaining of control and CHD7-knockdown cells after neuronal differentiation (at day 22) for βIII-tubulin, NeuN and smooth muscle actin (SMA) (middle panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. Quantification of the number of neuronal cells and smooth muscle cells is shown in the lower panel (n=3). Data are presented as the mean ± SEM (**p < 0.01, ***p < 0.001, ****p < 0.0001, two-tailed unpaired Student's t test).

4.3. CHD7-knockdown neuroepithelial cells acquire the transcriptional signatures of neural crest cells

To confirm whether CHD7 reduction leads to altered lineage identity, I generated and compared transcriptome profiles from control and CHD7-knockdown NE cells before and immediately after CHIR treatment (day 8 and day 13, respectively). I reasoned that transcriptional dysregulation of lineage-specific markers might be enhanced in the progression towards differentiation. Gene Ontology (GO) analysis showed that "neuron differentiation," "axonogenesis" and "neuron projection" were among the top biological functions downregulated in CHIR-treated sh*CHD7*-expressing cells, whereas "skeletal system development," "cell and biological adhesion" and "extracellular matrix" were among the main biological functions upregulated in these cells (Fig.7).

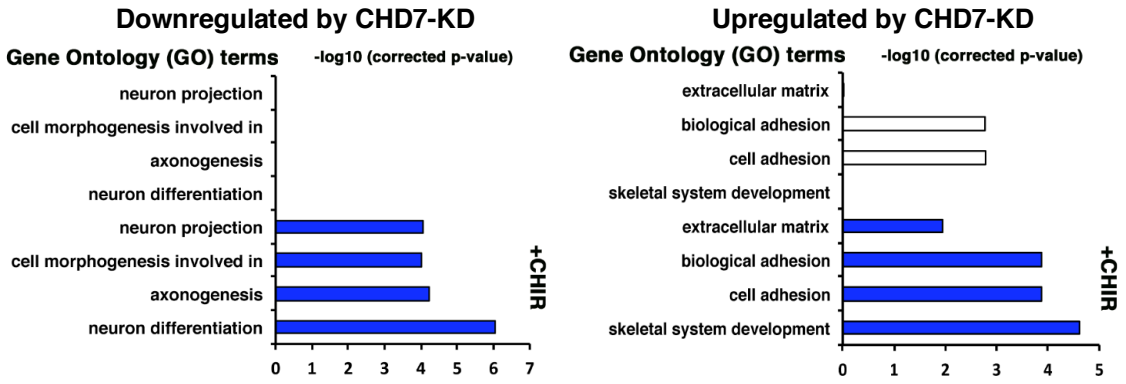


Figure 7: Gene expression programs dysregulated in CHD7-knockdown cells.
Bar plot of the top four GO terms for genes downregulated (left panel) and upregulated (right panel) by CHD7 knockdown in CHIR-treated NE cells (blue bars). Enrichment of corresponding GO terms in NE cells is also shown (white bars).

Notably, a subset of dysregulated NC signature genes was preferentially enriched in the GO term associated with "skeletal system development" (Simoes-Costa and Bronner,

2013), and I further validated the upregulation of some of these transcription factors (TFs), including *TWIST1* (Sauka-Spengler and Bronner-Fraser, 2008), *SNAI2* (Nieto et al., 1994; Thomas et al., 2008), and *MSX1/2* (Hill et al., 1989; Nikitina et al., 2008) by qRT-PCR analyses (Fig. 8). These results suggest that acquisition of NC-like identity occurs as a consequence of the loss of NE identity.

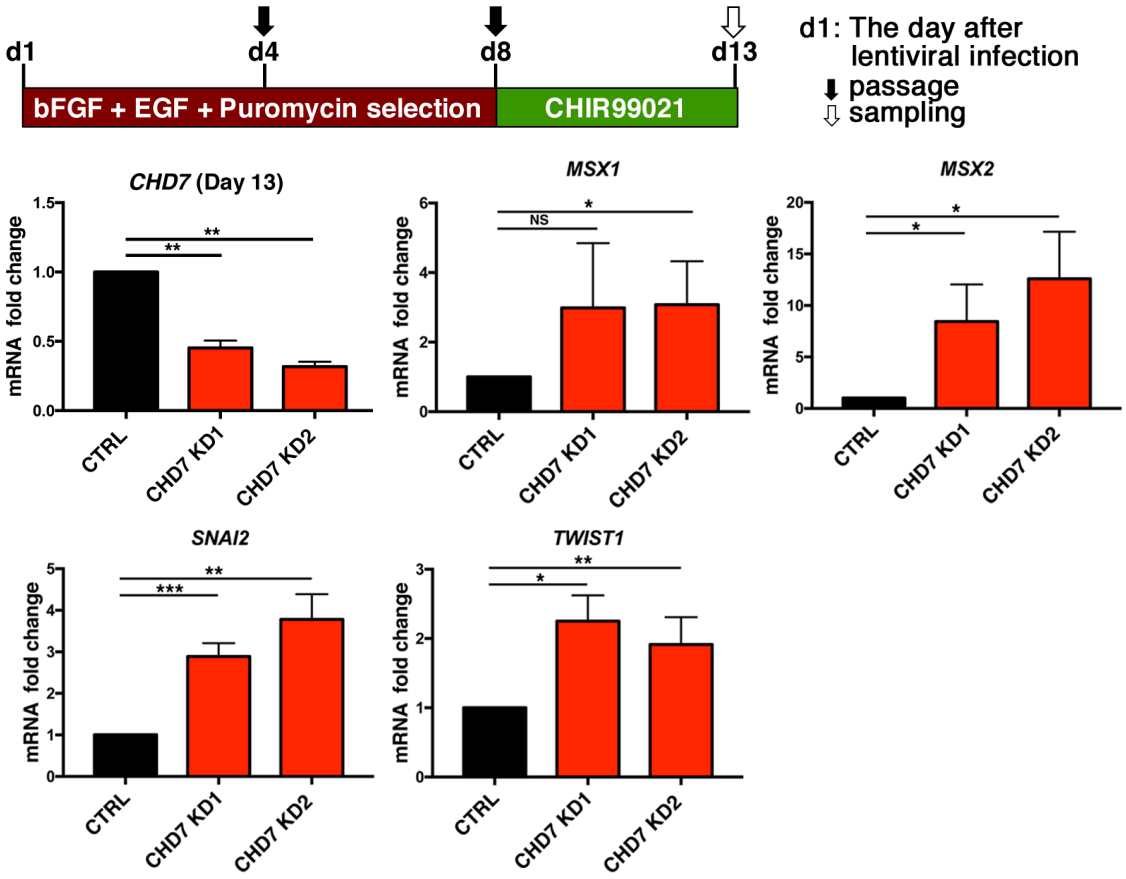


Figure 8: NC genes dysregulated in CHD7-knockdown cells.
Schematic diagram illustrating collection of RNA samples for analysis (upper panel). qRT-PCR validation of *CHD7*, *MSX1/2*, *SNAI2* and *TWIST1* in CHIR-treated control and CHD7-knockdown samples (lower panel) (n=3). Data are presented as the mean ± SEM (*p < 0.05, **p < 0.01 and ***p < 0.001, two-tailed unpaired Student's t test).

To further validate this finding, I analyzed whether the gene expression program evolved from NE specific to NC specific in CHD7-knockdown cells by using pairwise correlation of neural progenitor cell (NPC)-enriched and neural crest cell (NCC)-enriched genes extracted from publicly available data sets (Fig. 9). The global gene expression profiles of CHD7-knockdown cells were more similar to those of NCCs than those of NPCs; conversely, control knockdown cells remained closely correlated to NPCs (Fig. 9).

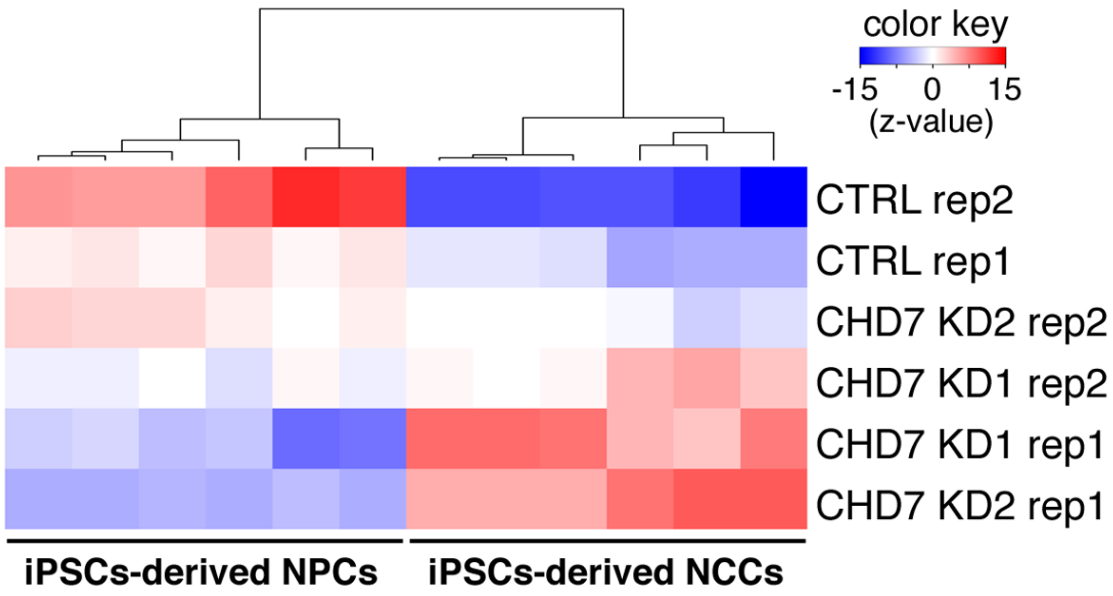


Figure 9: CHD7-knockdown NE cells acquired the transcriptional features of NC cells.
Heatmap representation of correlation of gene expression of CHIR-treated control and CHD7-knockdown NE cells with gene expression of iPSC-derived neural progenitor cells (NPCs) and iPSC-neural crest cells (NCCs). The color indicates the z-value for correlation significance.

Our current data suggest that at the end of CHIR treatment, the NC-specific gene expression program that reinforces NC identity was activated in CHD7-knockdown cells. To determine the earliest timing of transition into NC-like cells, I examined NE cells at 8 days (passage 1) and 12 days (passage 2) after transduction with lentiviruses expressing *CHD7* shRNA or control shRNA for the ectopic expression of SOX9 and SNAIL. Forced expression of these two TFs is sufficient to induce NC-like characteristics and promote epithelial-mesenchymal transition of NE cells, ultimately leading to inhibition of CNS neuronal generation (Cheung and Briscoe, 2003; Cheung et al., 2005). As shown in Fig. 10, immunocytochemistry and qRT-PCR analyses collectively showed that SOX9 and SNAI2 were upregulated in cells expressing *CHD7* shRNA as early as day 8 and persisted until day 12 after transduction in NE maintenance medium. Indeed, discernible changes in NE cell morphology that comprised the loss of rosette-like patterns and the acquisition of mesenchymal-like loosely packed structures were observed the day after the first passage (day 5) (data not shown). Together, our findings suggest that a high expression level of CHD7 is directly correlated with NE identity. Failure to maintain a robust CHD7 level triggered a rapid cell fate switch, as shown by the manifestation of transcriptional and morphological features of NCCs upon CHD7 knockdown. These data suggest that the loss of NE identity and the subsequent conversion of CNS stem cells to non-CNS cells likely cause prevalent terminal differentiation defects in $Chd7^{-/-}$ mouse neural progenitors *in vivo* (Feng et al., 2017; Feng et al., 2013; He et al., 2016).

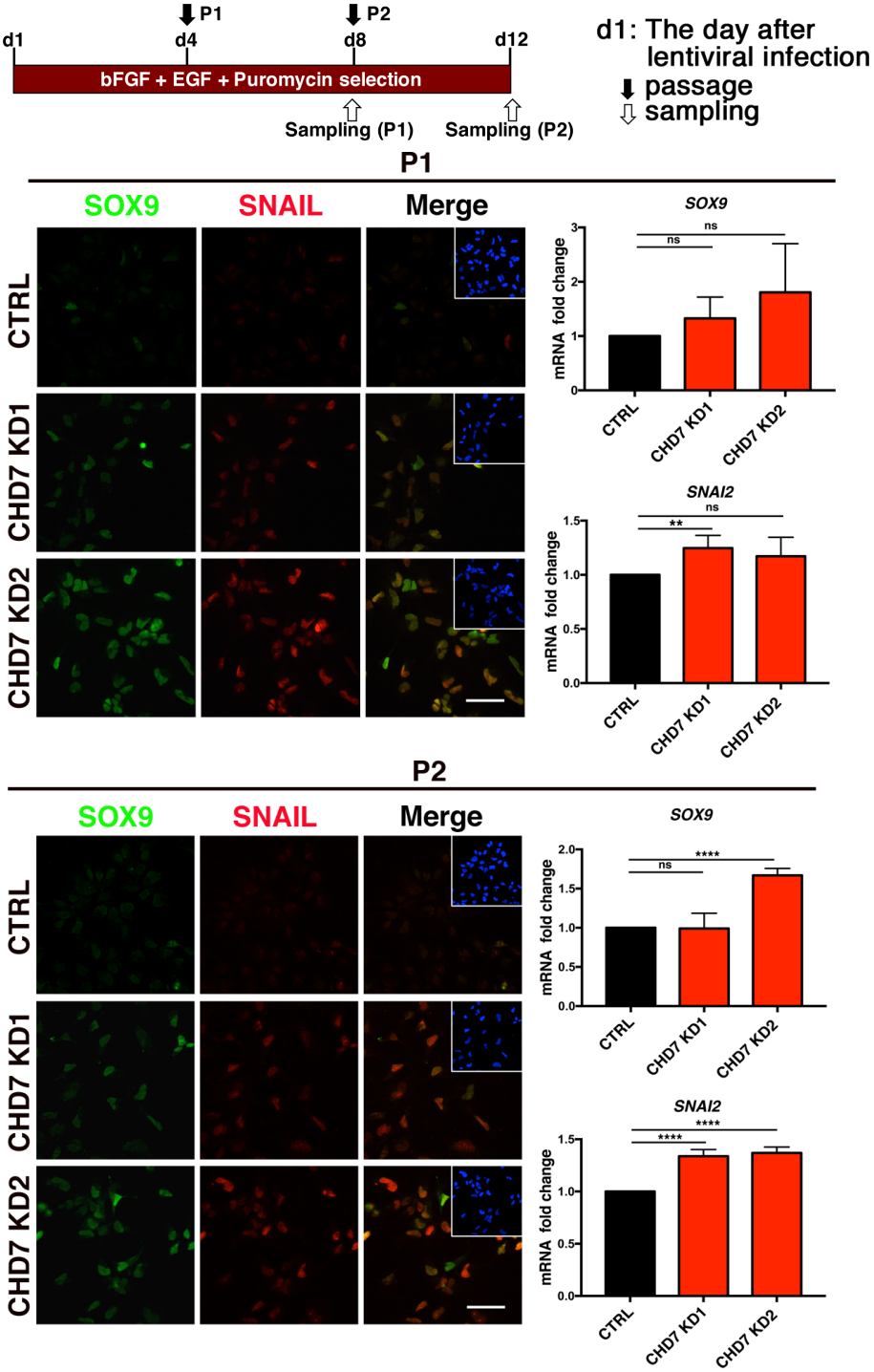


Figure 10: NC genes were dysregulated in CHD7-knockdown cells since day 8 after CHD7 knockdown.

Schematic diagram illustrating fixation and analysis of CHD7-knockdown cells (upper panel). Immunocytochemistry of control and CHD7-knockdown NE cells for SOX9 and SNAIL expression at day 8 (P1) (left-upper panel) and day 12 (P2) (left-lower panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. qRT-PCR analyses of *SOX9* and *SNAI2* expression in control and CHD7-knockdown NE cells at day 8 (P1) (right-upper panel) and day 12 (P2) (right-lower panel) (n=3). Data are presented as mean ± SEM (ns > 0.05, **p < 0.01, ****p < 0.0001, two-tailed unpaired Student's t test).

4.4. CHD7 haploinsufficiency underlies inappropriate lineage commitment

To confirm whether the above lineage defect is relevant in patient-derived NE cells, I obtained NE cells from iPSCs derived from two individual CHARGE patients (Okuno et al., 2017) and two unrelated healthy control individuals (hereafter referred to as wild-type) (Isoda et al., 2016a; Takahashi et al., 2007). The first CHARGE syndrome patient is heterozygous for CHD7^{Gln1391fs} (CHD7$^{G1391fs/+}$, hereafter referred to as CH1), and the second patient is heterozygous for CHD7^{Arg1494X} (CHD7$^{R1494X/+}$, hereafter referred to as CH2) (Fig. 11).

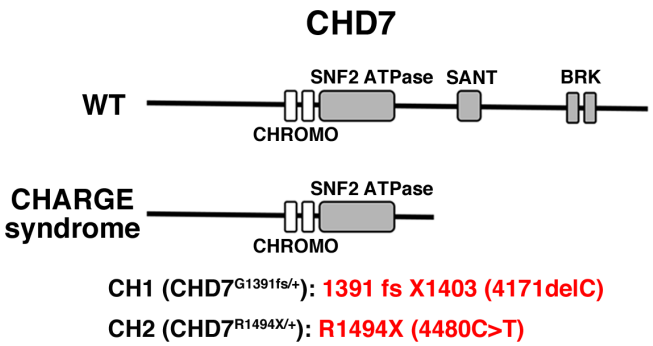


Figure 11: CHD7 mutations carried by CHARGE patient-derived cells.
Illustration of frameshift and nonsense mutations in *CHD7* harbored by the two CHARGE syndrome patients used in this study.

Both frameshift and nonsense mutations lead to premature termination of CHD7, which is predicted to undergo degradation by nonsense-mediated decay (NMD) (Zentner et al., 2010). Consistently, I found that *CHD7* mRNA and protein were reduced in patient-derived iPSCs compared with those in wild-type iPSCs (Fig. 12). To examine their

differentiation capacity, I generated NE cells from wild-type and patient-derived iPSCs (Fig. 13) and induced neuronal differentiation as described earlier.

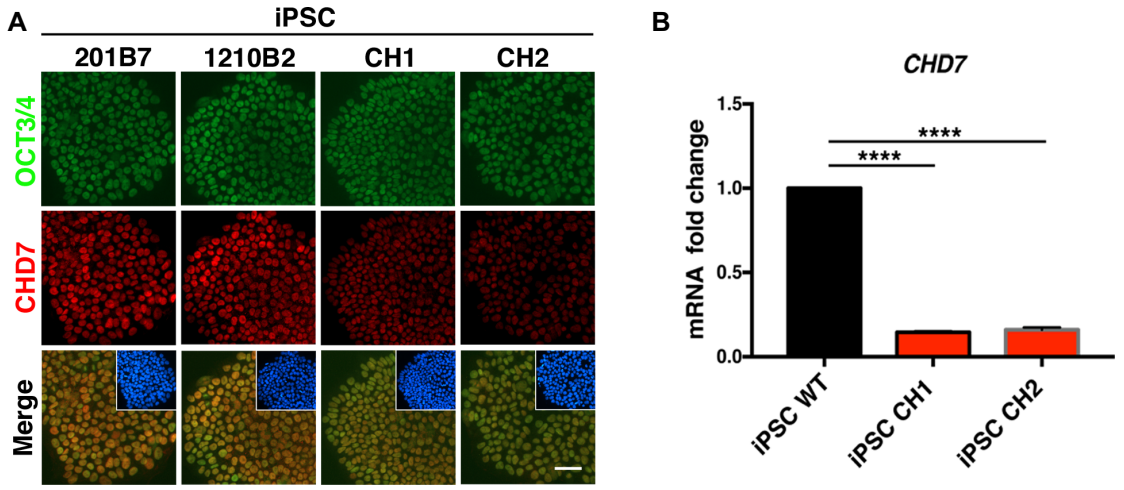


Figure 12: Reduced CHD7 expression in CHARGE patient-derived iPS cells.
(A) Immunocytochemistry for OCT3/4, a marker for pluripotent stem cells, and CHD7 in wild-type (201B7 and 1210B2) and CHARGE (CH1 and CH2)-iPS cells. Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. (B) qRT-PCR analyses of *CHD7* expression in wild-type and CHARGE-iPS cells (n=3). Data for wild-type cells depicts the average CHD7 expression in different clones of two wild-type iPSC lines (WD39 and 201B7). Data for CHARGE-iPS cells depicts the average CHD7 expression in six different clones of CH1 and seven different clones of CH2. Data are presented as the mean ± SEM (****p < 0.0001, two-tailed unpaired Student's t test).

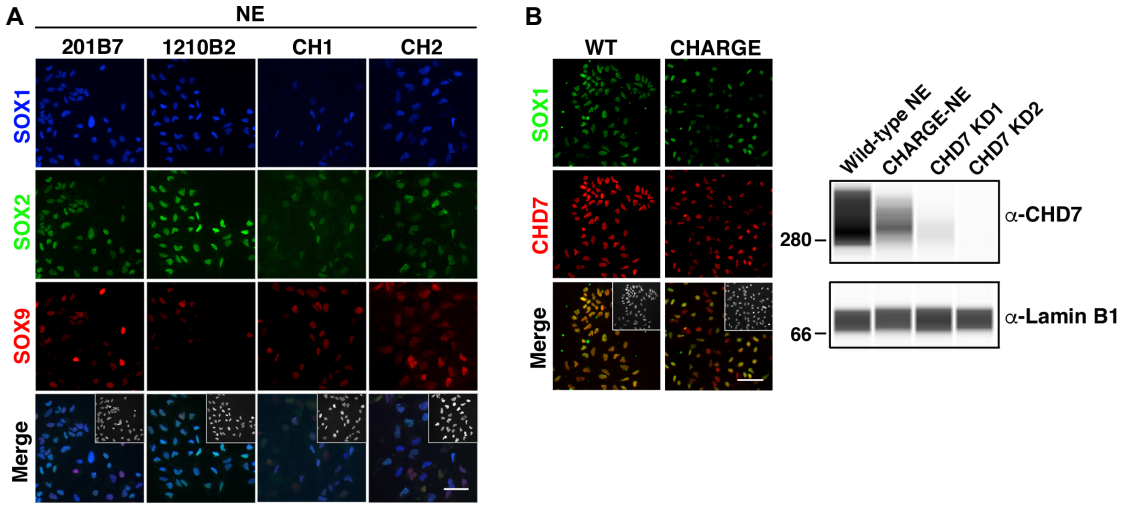


Figure 13: Reduced CHD7 expression in CHARGE patient derived iPS-NE cells.

(A) Immunofluorescence detection of SOX1 (blue), SOX2 (green) and SOX9 (red) in wild-type and CHARGE-NE cells. Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm.
(B) Immunocytochemistry for SOX1 and CHD7 in control (AF22) and CHARGE-NE cells. Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. Immunoblot analysis of CHD7 expression in control (AF22), CHARGE-NE (CH2) cells, and CHD7-knockdown NE cells. LaminB1 was used as a nuclear loading control (right panel).

Wild-type iPSC-NE cells were competent to differentiate into βIII-tubulin-positive neurons (Fig. 14). By contrast, CHARGE patient-derived iPSC-NE (CHARGE-NE) cells exhibited substantially reduced neuronal differentiation capacity and spontaneously gave rise to SMA-positive cells (Fig. 14).

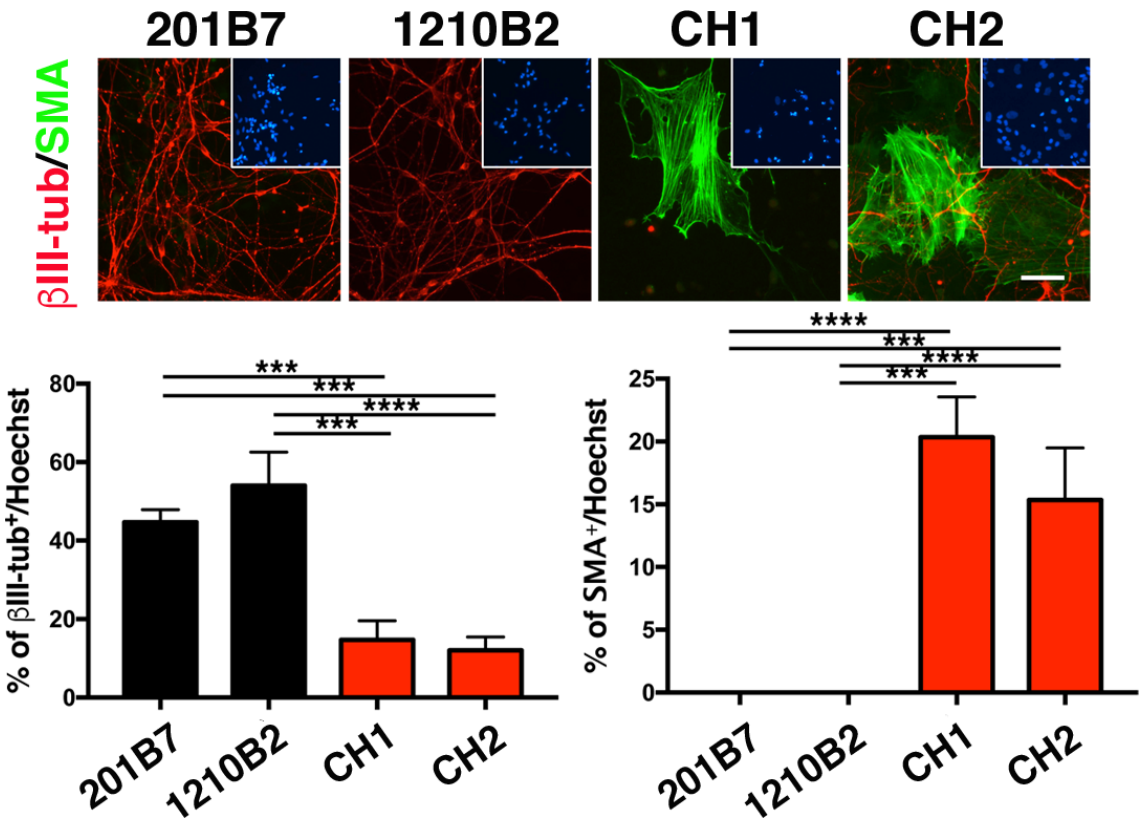


Figure 14: CHARGE patient-derived NE cells displayed altered differentiation propensity. Immunostaining of wild-type NE cells (201B7 and 1210B2) and CHARGE-NE cells (CH1 and CH2) after neuronal differentiation (at day 22) for βIII-tubulin (red) and SMA (green) (upper panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. Quantification is shown in the lower panel (n=3). Data are presented as the mean ± SEM (***$p < 0.001$ and ****$p < 0.0001$; one-way analysis of variance (ANOVA) with Bonferroni's post hoc test).

These findings suggest that CHD7 insufficiency contributes to the loss of neurogenic competence that causes the ensuing derepression of a non-neural differentiation program. Consistent with this notion, when I overexpressed full-length CHD7 in CHARGE-NE cells, neuronal differentiation was restored, and conversion into smooth muscle cells was efficiently repressed (Fig. 15).

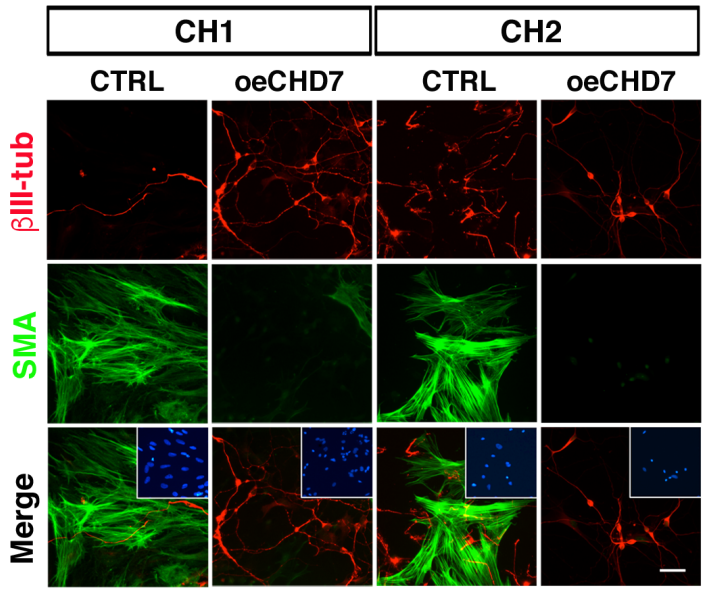


Figure 15: Forced expression of CHD7 restored neuronal differentiation ability.
Immunostaining of CHARGE-NE cells transfected with control vector (CTRL) or vector carrying full-length wild-type CHD7 (oeCHD7) for βIII-tubulin and SMA after neuronal differentiation (at day 22). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm.

Thus, our results suggest that defects in neuronal differentiation, a cellular program frequently affected by disruption of CHD7 expression, is caused by the aberrant transdifferentiation of NE cells to NC progenitors, leading to altered lineage output. I next examined whether the altered differentiation capacity of CHARGE-NE cells was associated with the acquisition of a NC transcriptional signature, which occurs in CHD7-

knockdown cells. By using the same sets of NPC- and NCC-enriched genes to compare correlations, I showed that the transcriptome profiles of CHARGE-NE cells were enriched for the NCC but not the NPC gene expression signature; conversely, wild-type NE cells corresponded to the NPC state (Fig. 16).

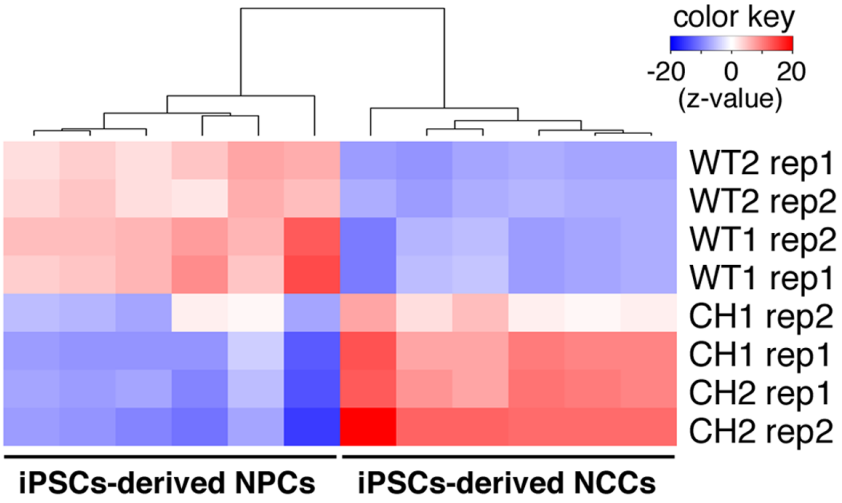


Figure 16: CHARGE patient-derived NE cells share transcriptional features of NC cells.
Heatmap representation of correlation of gene expression of wild-type and CHARGE-NE cells with gene expression of iPSC-derived NPC- and iPSC-NCC-enriched genes. The color indicates the z-value for correlation significance.

Approximately 80% of CHARGE syndrome patients harbor either a nonsense (44%) or frameshift mutation (34%) (Janssen et al., 2012); if these mutated transcripts escape NMD, they can be translated into truncated CHD7 protein. I asked whether CHD7^{G1391fs} and CHD7^{R1494X} could exert a dominant negative effect on wild-type CHD7. I confirmed the expression of mutant CHD7 proteins in HEK293T cells automated Western blotting (Fig. 17A). Next, I established NE cell lines stably expressing mutant CHD7 proteins and differentiated them into neurons. Neither the expression of CHD7^{G1391fs} nor CHD7^{R1494X} caused aberrant differentiation capacity; both cell lines generated βIII-tubulin-positive neurons with efficiencies comparable to those of wild-type NE cells, and no smooth muscle cells were observed (Fig. 17B).

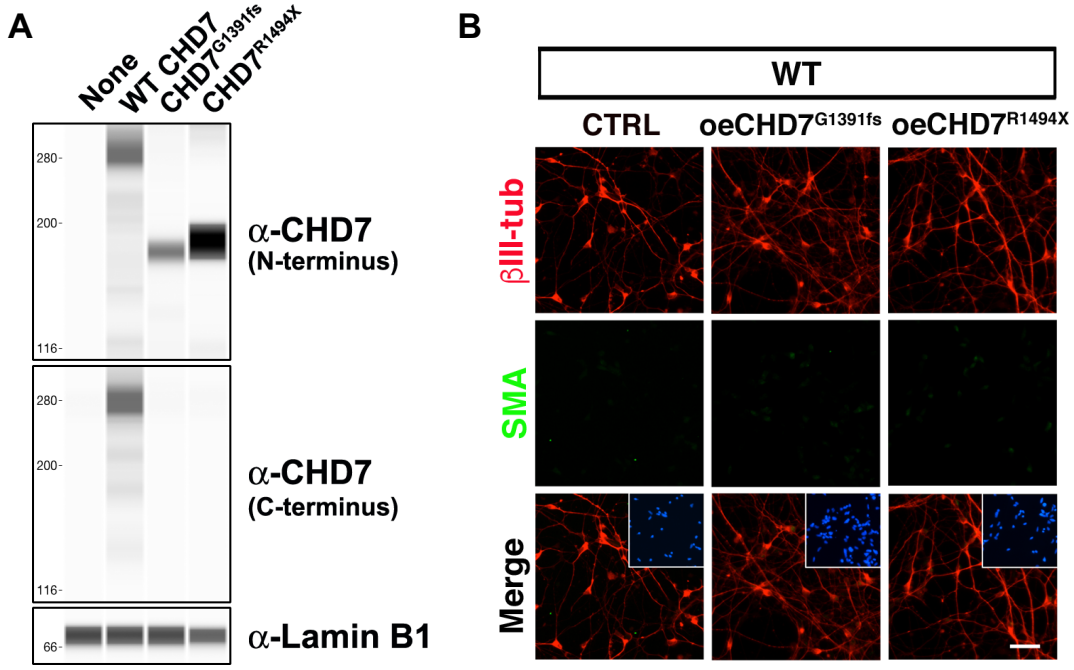


Figure 17: Unchanged neuronal differentiation ability in NE cells expressing mutant CHD7 proteins.
(A) Immunoblot analysis for CHD7 in HEK293T cells overexpressing full-length wild-type CHD7 and mutant CHD7 proteins (CHD7^{G1391fs} and CHD7^{R1494X}). Two different antibodies directed against N- or C-terminal of CHD7 were used. LaminB1 was used as a nuclear loading control.
(B) Immunocytochemistry showing βIII-tubulin and SMA expression for AF22 cells carrying control vector and AF22 cells carrying mutant CHD7 proteins (CHD7^{G1391fs} or CHD7^{R1494X}) after neuronal differentiation. Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm.

To ascertain the molecular identities of the respective NE cell lines that I established, I performed microarray analysis to characterize the global transcriptome profiles of wild-type NE cells, NE cells overexpressing $CHD7^{G1391fs}$ or $CHD7^{R1494X}$, and CHARGE-NE cells. Our clustering analyses revealed that wild-type NE cells and NE cells overexpressing $CHD7^{G1391fs}$ or $CHD7^{R1494X}$ clustered together in a group that was clearly separated from CHARGE-NE cells (Fig. 18). The expression of NPC-enriched genes, such as *SOX2*, *MSI1*, *NES*, *PAX6* and *PLZF*, remained high in wild-type NE cells and NE cells overexpressing $CHD7^{G1391fs}$ or $CHD7^{R1494X}$. Conversely, CHARGE-NE cells were characterized by aberrant upregulation of NCC-enriched genes, such as *SNAI1/2*, *MSX1/2*, *SOX9* and *TWIST1/2*, concomitant with the downregulation of *CHD7* (Fig. 18). Such distinction, together with our earlier observation that forced expression of CHD7 rescued the aberrant phenotype in CHARGE-NE cells, thus suggests that reduced CHD7 expression causes global gene expression changes that underlie the pathological mechanisms of both *CHD7* frameshift and nonsense mutations.

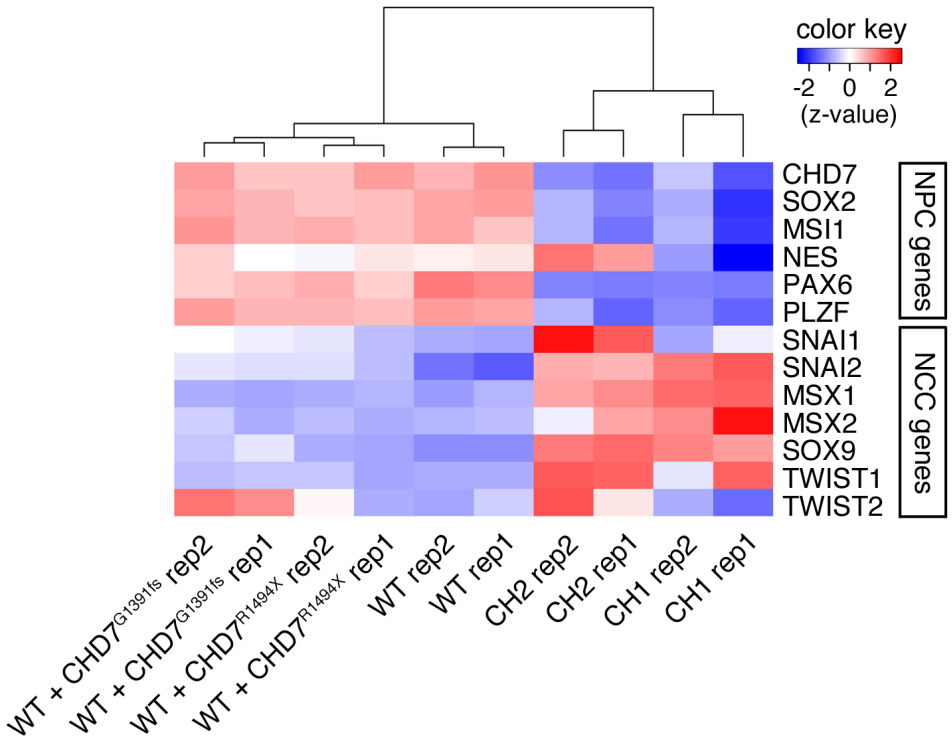


Figure 18: Dysregulated NSC and NC genes in CHARGE-NE cells
Heatmap displaying the expression levels of 6 NPC-enriched genes and 7 NCC-enriched genes in wild-type NE cells, NE cells expressing mutant CHD7 proteins and CHARGE-NE cells. Representative genes for each group are listed (right). The color indicates the z-value for correlation significance.

4.5. CHD7 binds to the majority of super-enhancers in NE cells

To gain a better understanding of the molecular mechanisms by which CHD7 regulates NE cell identity, I performed ChIP-seq (Johnson et al., 2007) to identify genomic targets of CHD7 in NE cells. First, I generated ChIP-seq data for CHD7 using two commercially available α-CHD7 antibodies (CHD7 ChIP-seq) but observed an apparent lack of ChIP-seq density overall (Fig. 19). I reasoned that single fixation of cells with formaldehyde might be inefficient, especially when it has not been clarified whether CHD7 binds directly to DNA. I then generated another set of ChIP-seq data for α-CHD7 (CST) using dual cross-linking reagents formaldehyde and ethylene glycol bis

(succinimidyl succinate) (EGS) (Yu et al., 2015) and identified 22,939 binding sites in NE cells (Fig. 19).

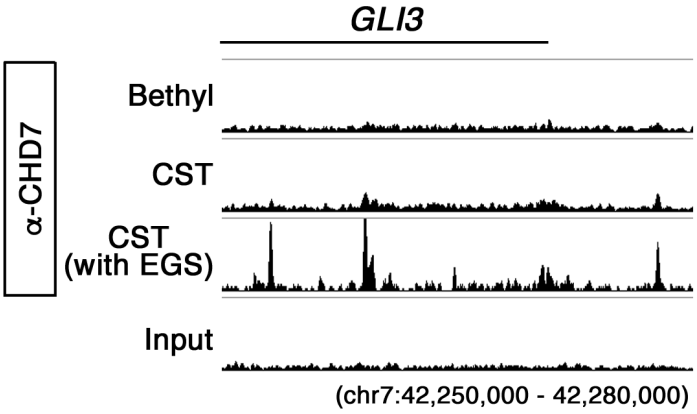


Figure 19: The identification of CHD7 binding targets.
Snapshot showing ChIP-seq signals for CHD7 at *GLI3* locus and comparison of signal densities for different antibodies and fixation conditions. The initial rounds of ChIP-seq experiments were performed using two commercially available α-CHD7 antibodies purchased from Bethyl (Bethyl) and Cell Signaling Technology (CST). ChIP-seq experiment performed using α-CHD7 from CST in cells fixed with formalin and EGS (with EGS, also referred to as (CST/EGS)) yielded higher signal intensity with sharper peaks.

Although the dual cross-linking protocol significantly improved the immunoprecipitation efficiency of α-CHD7 (CST), owing to its ability to capture indirect DNA-protein associations, it inevitably included false positive sites. To address this issue, I established two individual clones of NE cells (#CB1 and #F10) stably expressing a human influenza hemagglutinin (HA) tag integrated into the endogenous *CHD7* locus using CRISPR-Cas9-mediated genome editing (Fig. 20A-C) (Cong et al., 2013; Doudna and Charpentier, 2014; Hsu et al., 2014; Sander and Joung, 2014; Savic et al., 2015) and generated alternative ChIP-seq sets using α-HA (Fig. 21). Integration of α-CHD7 (CST)/EGS and α-HA ChIP-seq data sets identified 12,345 common binding sites, and 10,986 of these sites represented high-confidence CHD7-bound sites.

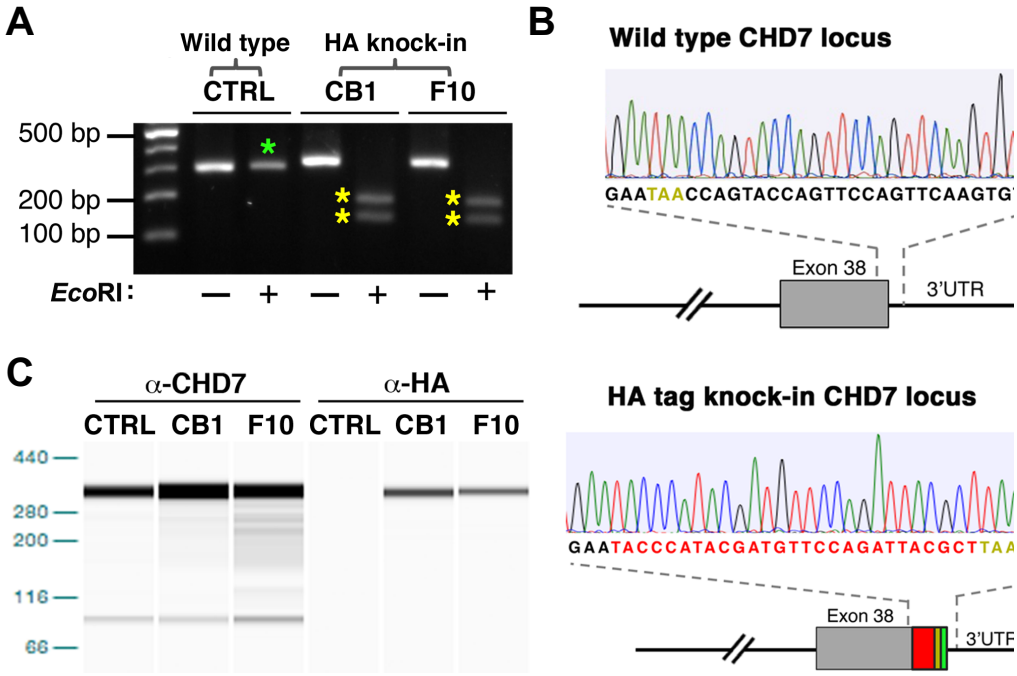


Figure 20: The generation of CRISPR/Cas9-mediated knockin CHD7-HA NE cell lines.
(A) Genotyping PCR followed by restriction enzyme treatment in control parental NE cells AF22 (CTRL), edited NE cells # CB1 and # F10. Green asterisk marks undigested band (302 bp), while yellow asterisks mark digested fragments at 193 bp and 139 bp.(B) Sequencing of *CHD7* locus in control and the edited clone. The red box represents the HA-tag, the yellow box indicates a stop codon and the green box denotes an *Eco*RI site. The coding sequence is in gray. Sequences for HA, stop codon and the *Eco*RI recognition motif are shown in red, yellow and green, respectively.(C) Immunoblot analysis of CHD7 and HA expression in AF22, #CB1 and #F10 cells. LaminB1 was used as nuclear loading control.

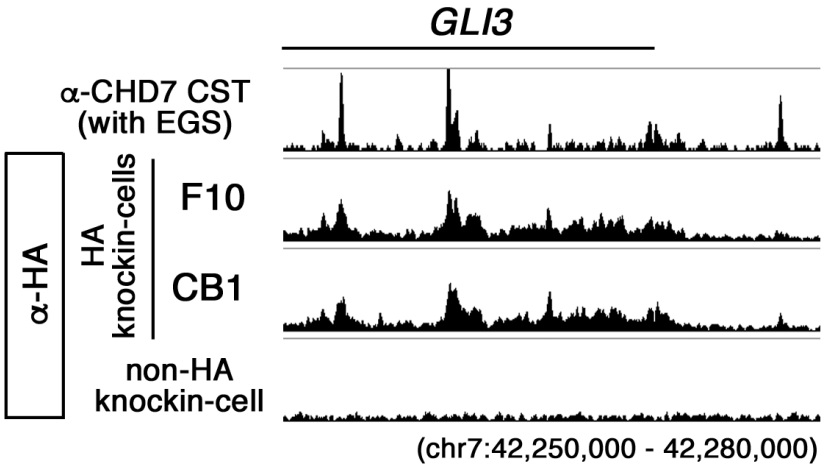


Figure 21: Significant overlap of binding sites between α-CHD7 ChIP-seq and α-HA ChIP-seq.
Snapshot showing ChIP-seq signals for CHD7 at *GLI3* locus. There is a high degree of overlap of CHD7 binding sites between α-CHD7 (CST/EGS) and α-HA ChIP-seq (F10 and CB1) data sets. ChIP-seq experiments using α-HA were performed in two independent HA knock-in cell lines, F10 and CB1. These cells were fixed with formaldehyde only.

GO analyses of CHD7-bound regions revealed the significant presence of GO terms related to neurulation, which includes neural tube formation, neural tube closure, and embryonic epithelial tube formation (Fig. 22A). This neurulation signature is strongly relevant to the hallmark features of CHARGE syndrome, given that the characteristic defects of this syndrome suggest dysfunctions in the neurulation process (Sanlaville and Verloes, 2007). Genome distribution analysis showed that 11% of CHD7 occupancy was located proximal to TSS, whereas 39% was located in intergenic regions (Fig. 22B).

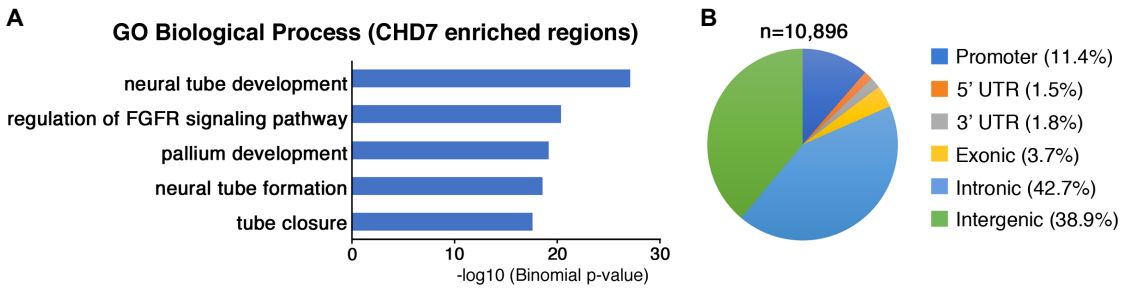


Figure 22: Genomic distribution and possible functions of CHD7 binding.
(A) GO analysis for biological processes associated with CHD7-bound regions.
(B) Genomic distribution of CHD7-bound regions in NE cells.

To gain further insights into the functional properties of CHD7, I generated chromatin landscapes of CHD7-bound regions, including H3K27ac, H3K4me1, H3K4me3, H3K9me3, H3K27me3, H3K36me3 and p300. Bioinformatics analyses revealed an overlap of regions bound by CHD7 with those of p300, H3K27ac, and H3K4me1 in which these marks are indicative of active enhancers (Fig. 23) (Creyghton et al., 2010). Conversely, CHD7 binding is most unlikely to occur within the promoter or an actively transcribed gene body region as evidenced by the minimal or lack of correlation with H3K4me3 and H3K36me3 (Fig. 23). These data confirmed the predominant localization of CHD7 to enhancer elements.

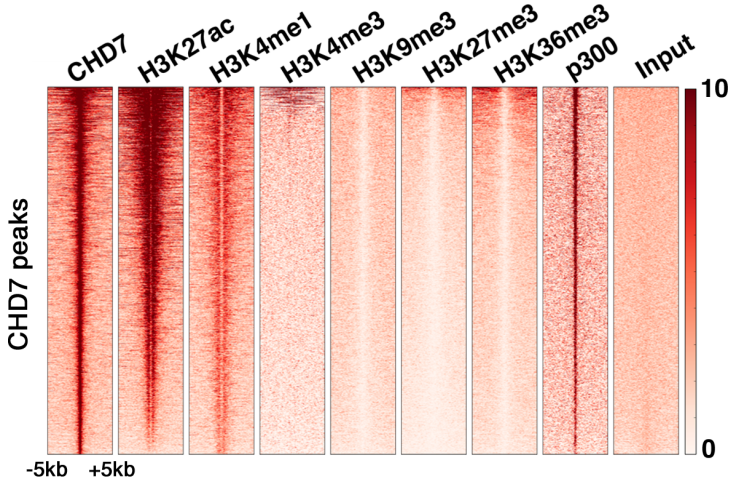


Figure 23: CHD7 selectively binds to active enhancers.
Heatmap analyses of ChIP-seq signals of CHD7 and selected histone modifications in NE cells. All ChIP-seq signals are displayed from ± 5 kb surrounding the center of each annotated CHD7 peak. Each CHD7 binding site is represented as a single horizontal line (red) centered on the peak summit.

A previous study has described Chd7 as one of the components enriched in super-enhancers (SEs) across the ESC genome (Hnisz et al., 2013). Given that SEs confer cellular identities (Adam et al., 2015; Whyte et al., 2013), I hypothesized that the maintenance of NE cell identity involves the binding of CHD7 to SEs. I first delineated

typical enhancers (TEs) and SEs in NE cells based on H3K27ac abundance using the ROSE algorithm, and identified 1,012 SEs (Fig. 24A) (Whyte et al., 2013). Although SEs constitute only 5% of active enhancers, I found that nearly all H3K27ac-defined SEs (92%) were CHD7 bound (Fig. 24B), suggesting that CHD7 is preferentially localized to SEs in human NE cells. Notably, SEs tend to be found proximal to genes encoding key regulators of neural development and neurogenesis, such as *ZIC4*, *BRN2*, *SOX21* and *SOX1* (Fig. 24A). Consistently, GO analysis of SE-associated genes revealed neural tube development as the enriched GO term (Fig. 24C), indicating that CHD7 localizes at SEs of NE cells to regulate neural/NE fate.

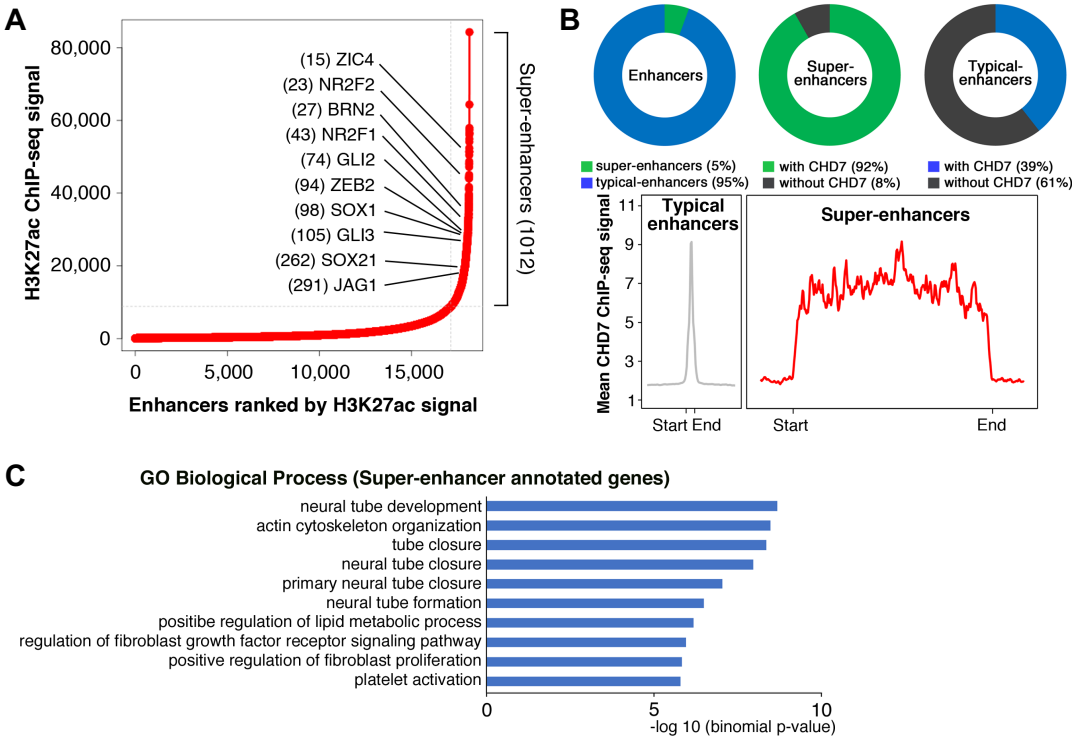


Figure 24: CHD7 selectively binds to super-enhancers.
(A) A population of enhancers that appears above the inflection point of the curve is considered super-enhancers. The respective ranks of super-enhancers and their associated genes are shown.
(B) Distribution of typical enhancers and super-enhancers in NE cells (upper left). Distribution of CHD7 binding that occurred at super-enhancers (upper middle) and typical enhancers (upper right). Metagene representations of mean ChIP-seq signal for CHD7 across typical enhancers and super-enhancer domains are shown in the lower panel.
(C) Top GO biological functions enriched among super-enhancers-associated genes.

4.6. CHD7 activates lineage-specific enhancers in human NE cells

To determine the enhancer properties of CHD7-bound distal regions, I evaluated these regions using the VISTA Enhancer Browser (Visel et al., 2007), a database of tissue-specific human enhancers with validated activity in mice. CHD7 binding sites were most enriched for CNS active enhancer elements (Fig. 25), including the neural tube, midbrain, hindbrain, and forebrain.

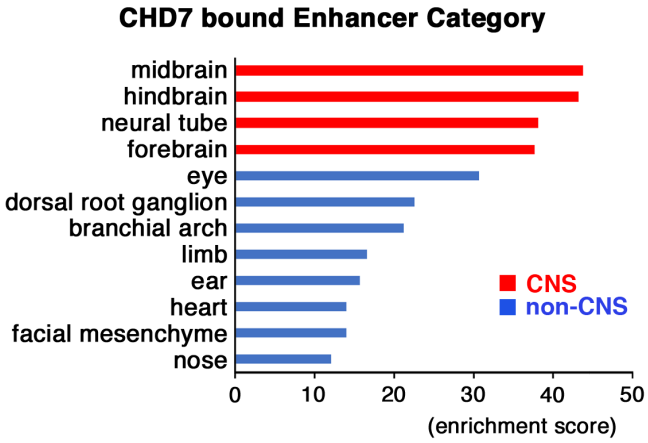


Figure 25: CHD7 selectively binds to CNS-specific enhancers in NE cells. Functional annotation of genomic regions bound by CHD7 intersected with enhancer regions validated in the VISTA Enhancer Browser database. Bar plot showing enrichment of CHD7 occupancy across 12 different tissues.

As a case in point, I found that CHD7 binding at SE-associated genes in NE cells, including *GLI3* and *ZIC1/4* loci, coincided with the presence of annotated enhancers that exhibit strong activity in the brain and neural tube (Fig. 26). Conversely, annotated enhancers without CHD7 either showed strong expression in limbs or were functionally inactive (Fig. 26). The identification of CNS-specific CHD7-enriched enhancers suggests that CHD7 regulates CNS-specificity of its corresponding genes and more importantly,

indicates a role for CHD7 in CNS lineage determination.

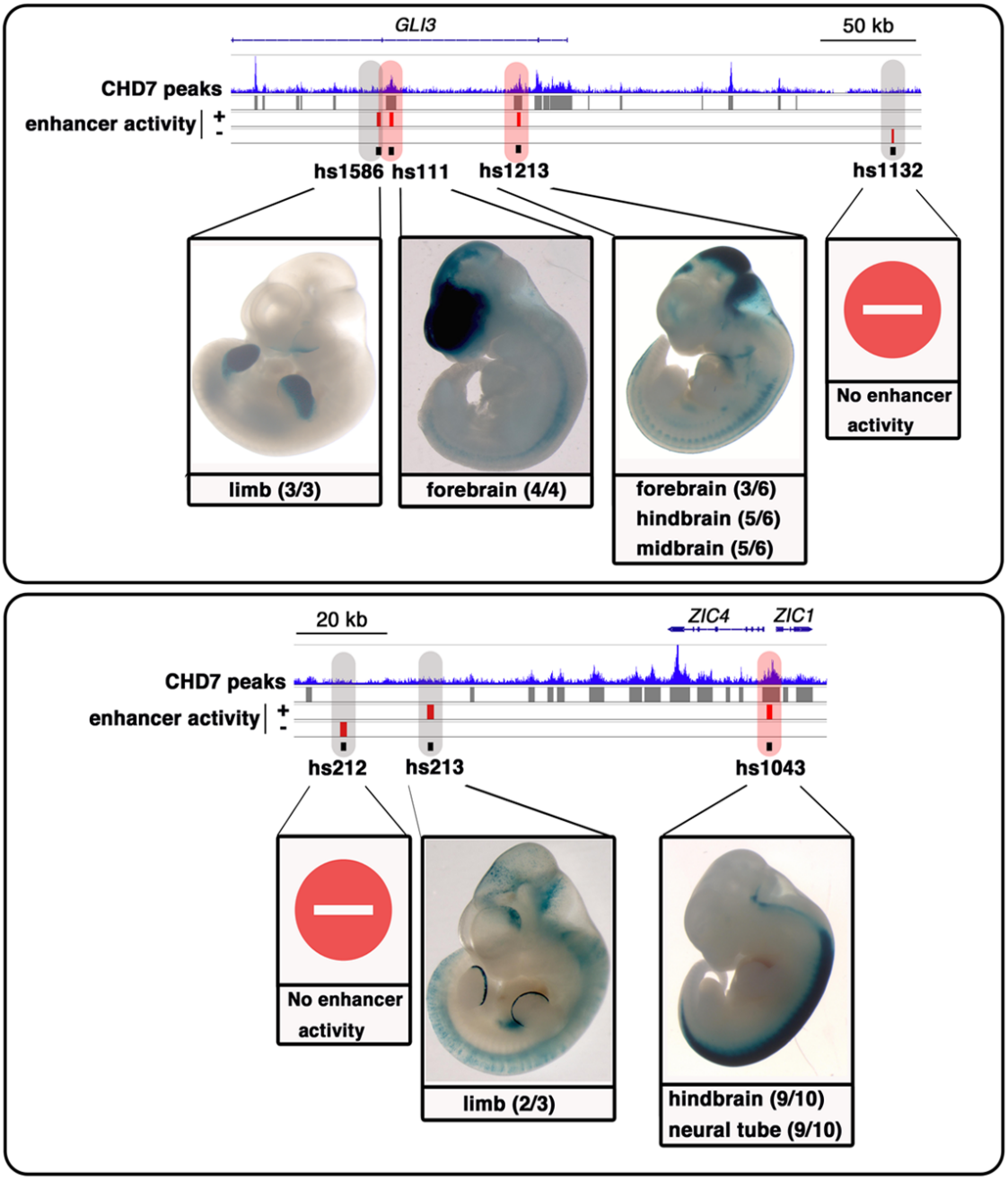


Figure 26: CHD7 selectively binds to CNS-specific enhancers in NE cells.
Snapshots showing combined tracks of CHD7 peaks (gray bars, extracted from α-HA ChIP-seq data) and VISTA-validated enhancer regions (red bars) near *GLI3* (up) and *ZIC1/4* (below). Shown below the tracks are screen shots of tissue-specific activities of the enhancers located proximal to *ZIC1/4* and *GLI3* tested in a lacZ reporter transgenic mouse assay and extracted from the VISTA Enhancer Browser database.

I found that CHD7 is selectively associated with CNS-specific enhancers, but little is known about CHD7-mediated enhancer regulation. In light of the notion that enhancer activity is correlated with the abundance of H3K27ac, I evaluated whether the reduction of CHD7 contributes directly to altered H3K27ac enrichment at CHD7 binding sites across the genome in CHD7-knockdown and CHARGE-NE cells. Compared with control-knockdown or wild-type NE cells, CHD7-knockdown NE cells and CHARGE-NE cells displayed an average decrease in H3K27ac at CHD7-bound regions (Fig. 27).

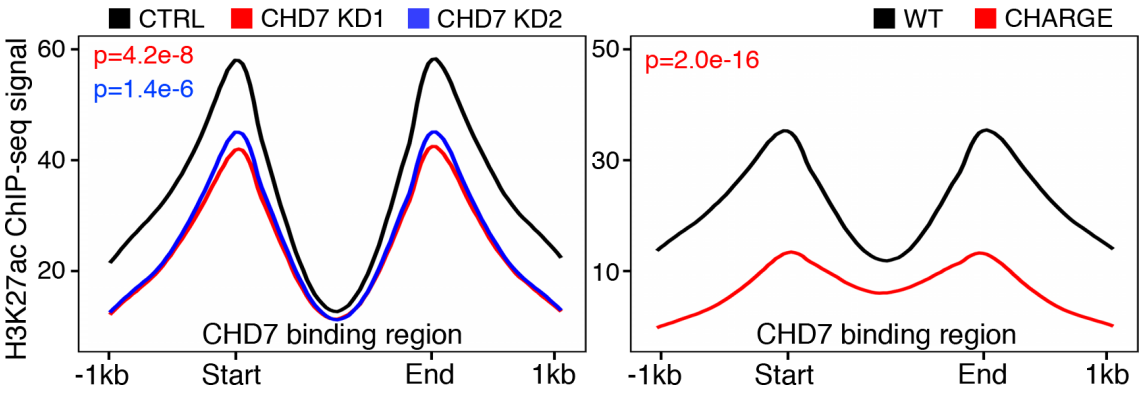


Figure 27: CHD7 is an epigenetic regulator in NE cells.
Density plots of mean H3K27ac ChIP-seq signals within ± 1 kb of CHD7 peak summits in control and CHD7-knockdown NE cells (left panel). Density plots of mean H3K27ac ChIP-seq signals within ± 1 kb of CHD7 peak summits in wild-type and CHARGE-NE cells (right panel).

I next determined whether H3K27ac surrounding SE regions was dependent on CHD7 abundance. Remarkably, I detected lower accumulation of H3K27ac in CHD7 haploinsufficiency than in wild-type, consistent with that of CHD7 knockdown (Fig. 28). Furthermore, I found that 30% of SE-associated genes were downregulated in both CHD7-knockdown and CHARGE-NE cells (but not of those upregulated; < 5%) (Fig.

29), establishing a role for CHD7 in regulating the activation of cell type-specific SEs and thus cell type-specific gene expression.

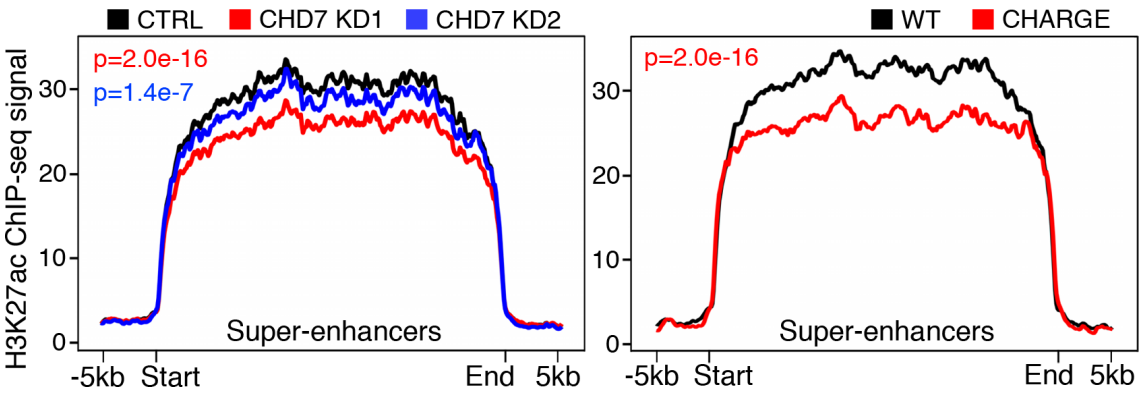


Figure 28: CHD7 is directly responsible for the maintenance of H3K27ac density at SEs.
Density plots of mean H3K27ac ChIP-seq signals encompassing super-enhancer domains in control and CHD7-knockdown NE cells (left panel). Density plots of mean H3K27ac ChIP-seq signals encompassing super-enhancer domains in wild-type and CHARGE-NE cells (right panel).

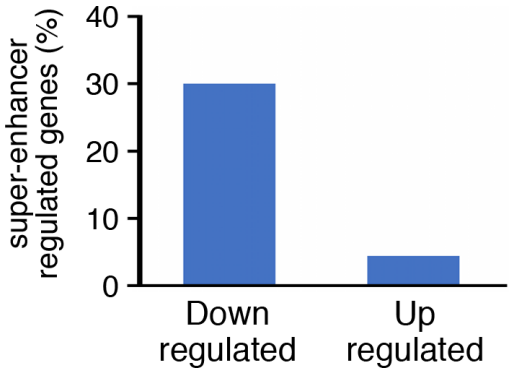


Figure 29: CHD7 regulates the activation of SE-associated genes.
(C) Bar plot showing the percentage of super-enhancer-associated genes, which are commonly downregulated or upregulated in CHD7-knockdown and CHARGE-NE cells.

I further analyzed whether CHD7 also modulates H3K27ac enrichment at CNS-specific enhancers identified using VISTA Enhancer Browser. I observed a decreased H3K27ac signal at CNS-specific enhancers of CHD7-knockdown cells, albeit to a lower degree than that of CHARGE-NE cells (Fig. 30). These results indicate a role for CNS-specific CHD7-enriched enhancer activity in the pathogenesis of CHARGE syndrome.

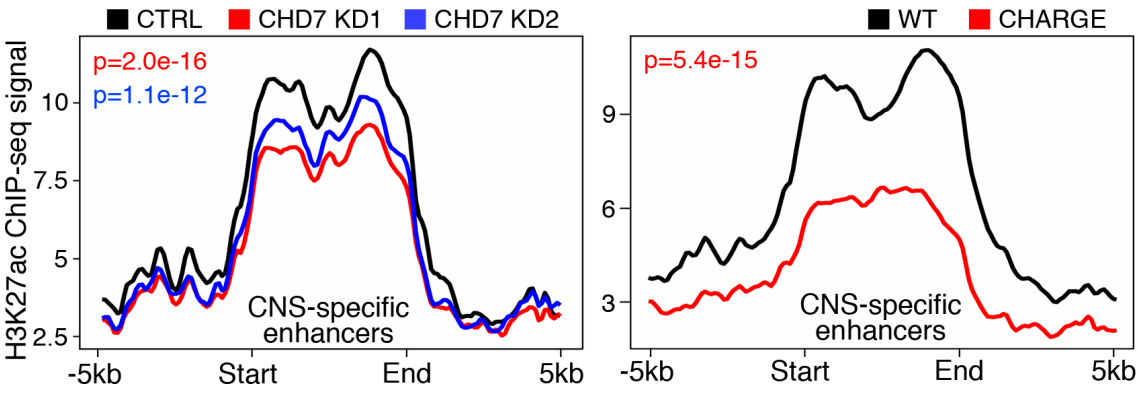


Figure 30: CHD7 maintains the H3K27ac profiles at CNS-specific enhancers in NE cells. Density plots of mean H3K27ac ChIP-seq signals at CNS-specific enhancers in control and CHD7-knockdown NE cells (left panel). Density plots of mean H3K27ac ChIP-seq signals at CNS-specific enhancers in wild-type and CHARGE-NE cells (right panel).

To support the above findings, I categorized dysregulated H3K27ac-occupied regions in CHD7-knockdown and CHARGE-NE cells based on the tissue-specific enhancer activity of the VISTA enhancer browser. The majority of CHD7-bound enhancers had a reduced H3K27ac signal in both CHARGE-NE and CHD7-knockcdown NE cells (Fig. 31). Furthermore, such enrichment was highly dependent on CHD7 occupancy as decreased H3K27ac regions that were not occupied by CHD7 were evenly distributed across different tissue-specific enhancers (Fig. 31). Nevertheless, non-CNS enhancers were overrepresented among CHD7-independent increased H3K27ac regions, indicating that

CHD7 is not directly responsible for the activation of a non-CNS fate program (Fig. 31). Taken together, our findings suggest that CHD7 is responsible for maintaining H3K27ac levels and tissue-specific enhancer activity; however, reduced expression of CHD7 in NE cells is insufficient to preserve this activation, ultimately leading to the disruption of cell type-specific gene expression and the CNS lineage program.

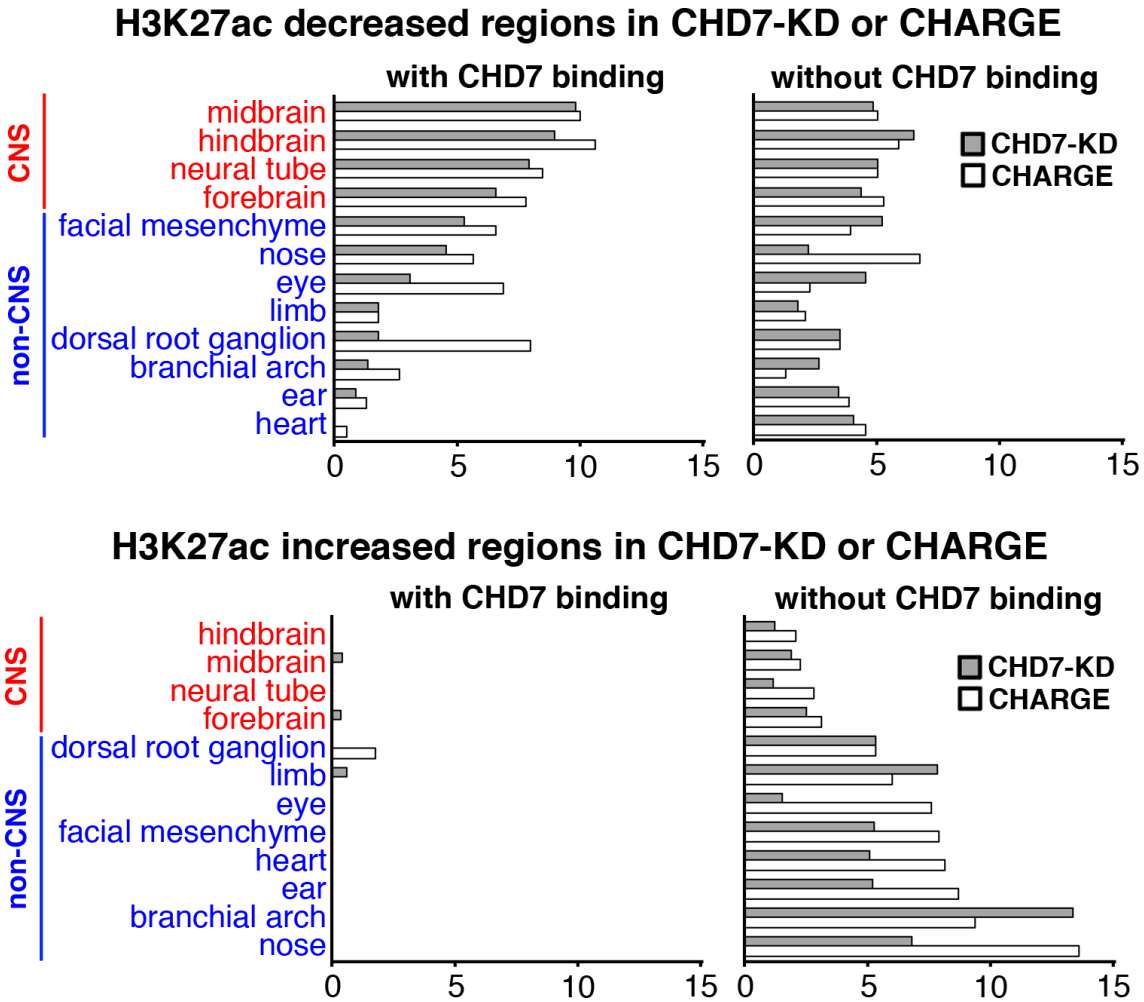


Figure 31: Dysregulated H3K27ac profiles in CHD7-knockdown and CHARGE-NE cells. Common CHD7-bound regions with decreased/increased H3K27ac levels identified in CHD7-knockdown and CHARGE-NE cells were intersected with enhancer regions validated in the VISTA Enhancer Browser database, and classified based on tissue-specific activity. These regions were further classified as with or without CHD7 binding.

4.7. CHD7-driven regulatory program in CHARGE pathogenesis

Having identified a role for CHD7 in controlling NE cell identity and CNS lineage determination, I next sought to identify CHD7-driven downstream effectors responsible for this pathway. I overlaid the gene list obtained from ChIP-Seq with the list of downregulated genes identified from the global gene expression profiles of control versus CHD7-knockdown and CHARGE-NE cells and identified 22 genuine candidates (Fig. 32A). I focused on downregulated genes given the association of CHD7 with active H3K27ac marks, an indication that CHD7 predominantly functions as a transcriptional activator. Two candidates, *SOX21* and *BRN2*, were of particular interest considering their roles in neurogenesis (Sandberg et al., 2005; Vierbuchen et al., 2010). To determine whether CHD7 directly activates *SOX21* and *BRN2* in NE cells, I examined the expression of *SOX21* and *BRN2* and found that the expression of these genes was reduced in CHD7-knockdown NE cells (Fig. 32B). In addition, I found that decreased *SOX21* and *BRN2* expression was accompanied by visible reductions in H3K27ac enrichment at CHD7 peaks proximal to *SOX21* and *BRN2* (Fig. 33A, B), further strengthening our conclusion that CHD7 epigenetically regulates its target genes.

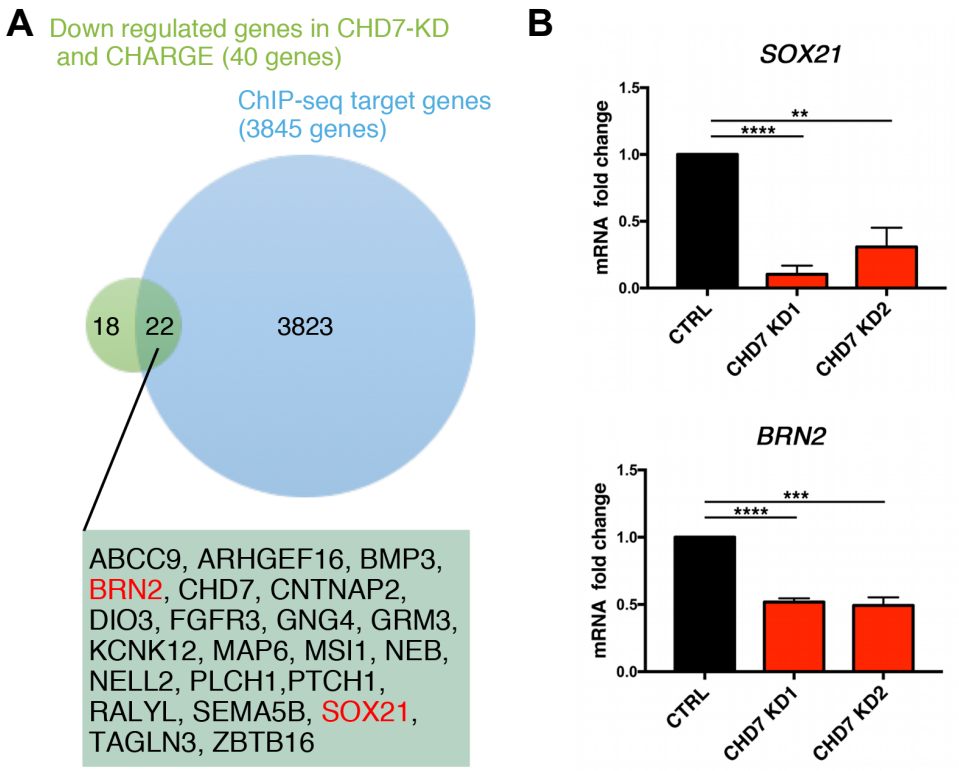


Figure 32: CHD7 regulates *SOX21* and *BRN2* expression.
(A) Venn diagram showing overlap between CHD7 genomic targets (3845 genes) and commonly downregulated targets in CHD7-knockdown and CHARGE-NE cells (40 genes). Twenty-two genes were identified as the cognate targets of CHD7.
(B) qRT-PCR validation of *SOX21* and *BRN2* in control and CHD7-knockdown samples (n=3). Data are presented as the mean ± SEM (**p < 0.01, ***p < 0.001 and ****p < 0.0001; two-tailed unpaired Student's t test).

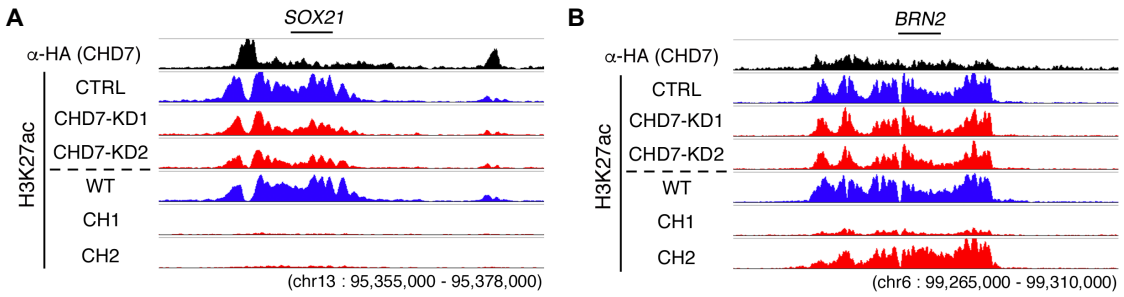
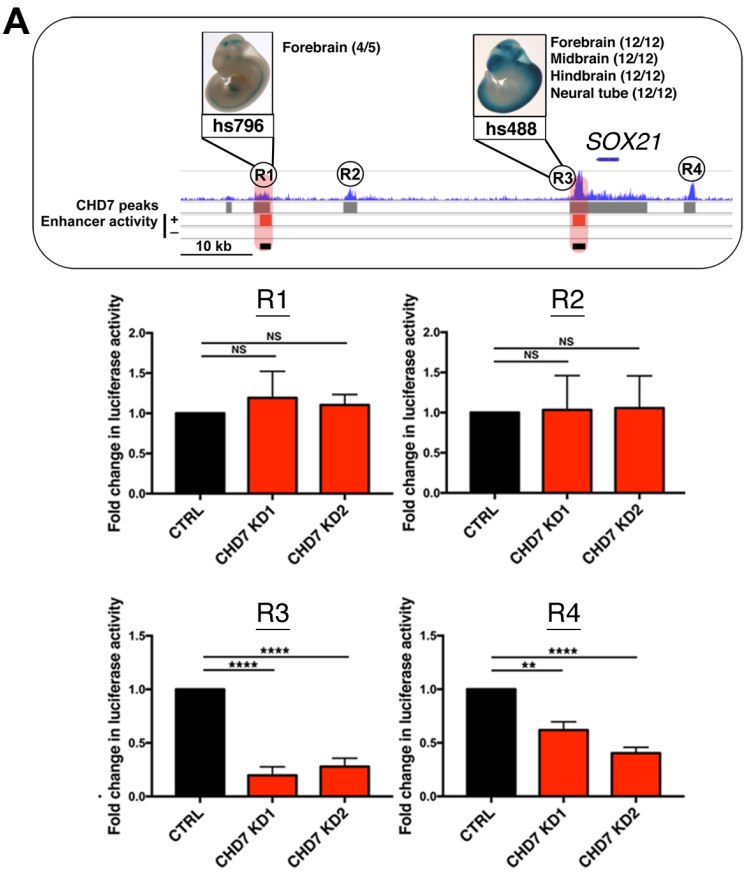


Figure 33: CHD7 regulates *SOX21* and *BRN2* expression by direct activation of super-enhancers.
CHD7 occupancy at the *SOX21* (A) and *BRN2* (B) genomic loci, together with the enrichment of H3K27ac at these loci in control-knockdown and wild-type NE cells versus CHD7-knockdown NE cells and CHARGE-NE cells. The ChIP-seq signal obtained using α-HA (HA knock-in into the *CHD7* locus) was used to depict CHD7 occupancy at these regions.

To further validate the reduction in H3K27ac levels around these regions, I examined the enhancer activities of several CHD7-bound enhancer regions surrounding *SOX21* and *BRN2*; I refer to these regions as R1–R4 (Fig. 34A, B). A search of the VISTA Enhancer Browser database revealed that three regions—R1 and R3 of *SOX21* and R1 of *BRN2*—are functionally validated enhancers with activities enriched in the CNS (Fig. 34A, B). The remaining regions are putative enhancers identified in this study. I found that five of eight regions tested were active in reporter assays (Fig. 35). Furthermore, the activities of four out of five active reporters (two per candidate gene) were clearly reduced in the presence of *CHD7* shRNA (Fig. 34A, B), suggesting that CHD7 regulates the gene expression of *SOX21* and *BRN2* through enhancer regions.



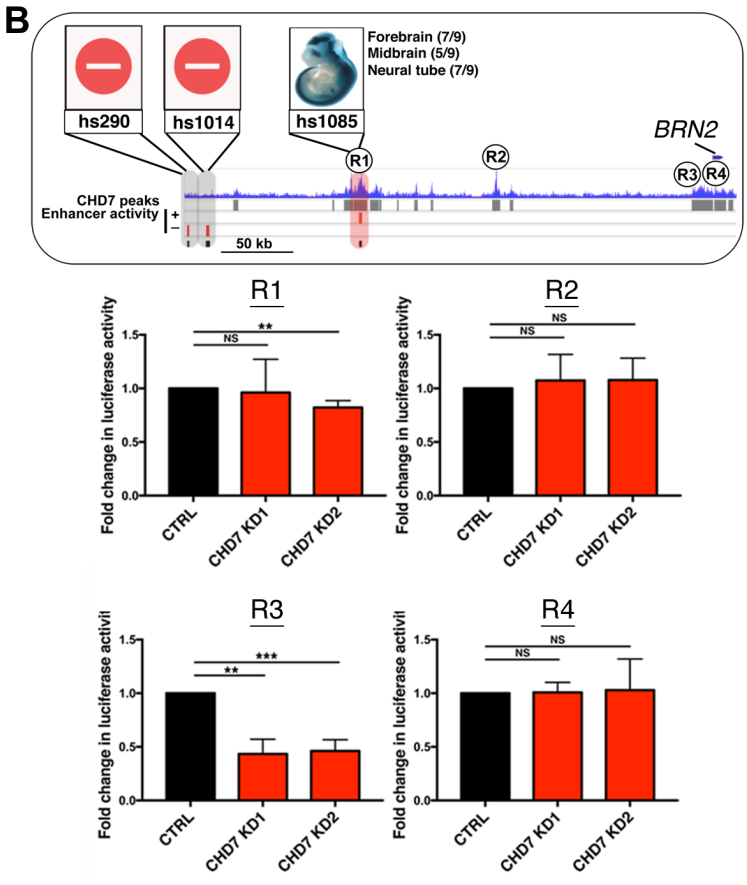


Figure 34: CHD7 regulates *SOX21* and *BRN2* expression by modulating enhancer acitivities. Representative tracks showing CHD7 peaks (gray bars) and VISTA-validated enhancer regions (red bars) around *SOX21* (A) and *BRN2* (B). Enhancer activities of distal genomic regions proximal to *SOX21* (hs488, R3 and hs796, R1) and *BRN2* (hs290, hs1014 and hs1085, R1) were validated using a lacZ reporter transgenic mouse assay and are available in the VISTA Enhancer Browser database. CHD7-bound regions with enhancer activities in the CNS are shaded in red (hs796 and hs488 for *SOX21* and hs1085 for *BRN2*). The regions with no recorded enhancer activity are shaded in gray (hs290 and hs1014). Data are presented as the mean ± SEM (**p < 0.01, ***p < 0.001 and ****p < 0.0001; two-tailed unpaired Student's t test).

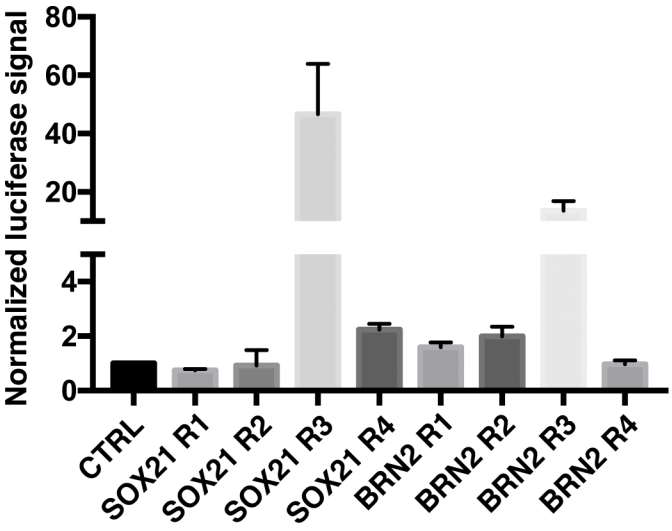


Figure 35: Basal activities of enhancer regions tested for luciferase reporter assay.
A luciferase reporter assay was performed in control and CHD7-knockdown NE cells to determine the activity of CHD7-bound enhancers near *SOX21* and *BRN2* (lower panel). A Renilla reporter was included for normalization (F/R ratio). Corresponding regions examined are shown in the upper panel.

Earlier, I showed that AF22 cells expressing $CHD7^{G1391fs}$ or $CHD7^{R1494X}$ exhibited no significant alterations in their differentiation ability, and no dominant negative effects were observed. Thus, I sought to further confirm that these mutants do not functionally inhibit CHD7-mediated enhancer activation. I examined the enhancer activities of two regions, namely, *SOX21*-R3 and *BRN2*-R3 in NE cells expressing mutant CHD7 proteins, and found that the inhibitory effects, if any, of mutant CHD7s are likely marginal (Fig. 36). These findings thus revealed the exquisite dosage requirement for CHD7 in preserving its activator function in gene regulation.

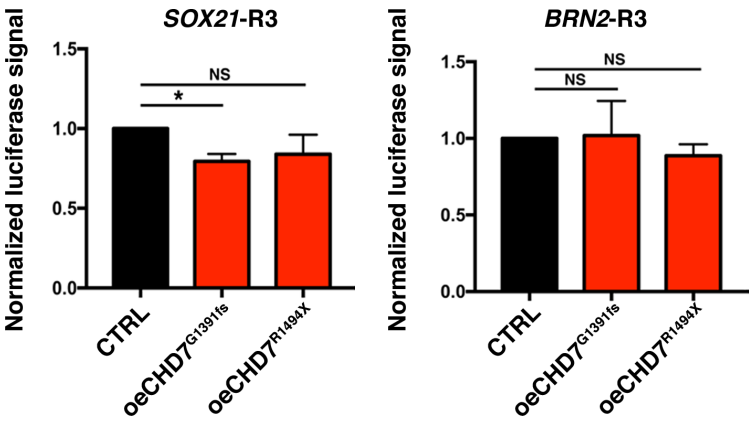


Figure 36: Mutant CHD7 proteins exert minimal dominant negative effects on wild-type CHD7. Luciferase reporter assay shows that overexpression of CHD7 mutant proteins in the NE cells hardly affect the activity of reporters carrying CHD7-bound enhancers near *SOX21* and *BRN2* (*SOX21*-R3 and *BRN2*-R3). Renilla reporter was included for normalization (F/R ratio). Corresponding regions examined were shown in the upper panel. Data are presented as the mean ± SEM (ns > 0.05 and *p < 0.05; two-tailed unpaired Student's t test).

To further confirm that CHD7 localization to these regions is NE specific, I compared CHD7 binding proximal to *SOX21* and *BRN2* with that in iPSCs and NCCs. I found distinguishable CHD7 peaks at genomic regions surrounding *SOX21* and *BRN2*, most strongly in the NE state compared with that in iPSC and NCC states (Fig. 37). Attesting to its cell type-specific binding pattern, CHD7 occupancy was selectively enriched in genomic regions proximal to *OCT4* and *NANOG* in the iPSC state, whereas distinct CHD7 peaks were detected proximal to *SOX9* and *SNAI1* in the NCC state (Fig. 37). These results support a cell type-specific regulatory role for CHD7.

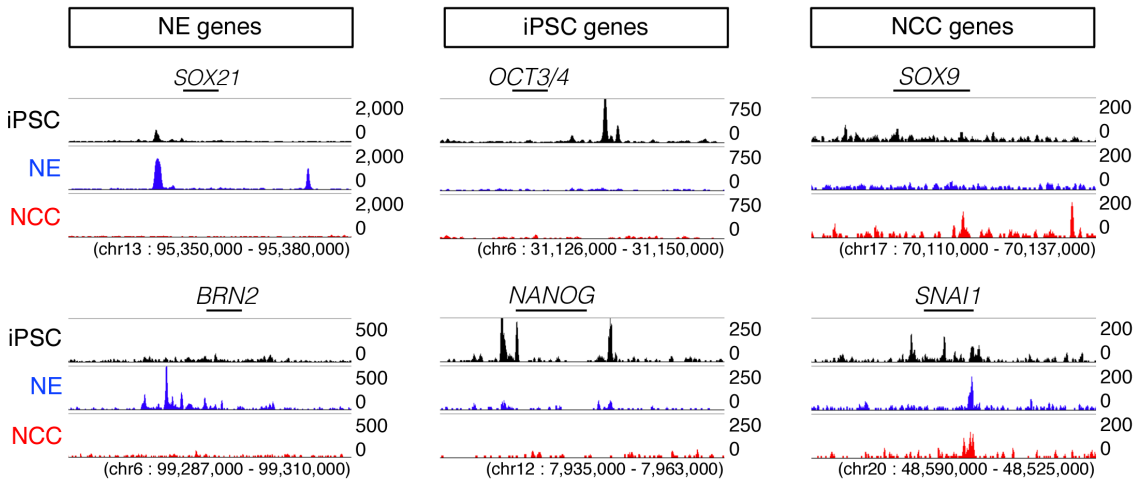


Figure 37: CHD7 plays cell type-specific role by binding to different sets of enhancers.
ChIP-seq signals for CHD7 at genomic loci specific to human NE cells (*SOX21* and *BRN2*) (left panel), human iPSCs (*OCT4* and *NANOG*) (middle panel) and human NCCs (*SOX9* and *SNAI1*) in each of these cell lines. The ChIP-seq signal for CHD7 (α-CHD7 (CST)/EGS) was used to depict CHD7 occupancy.

To further understand the role of SOX21 and BRN2 in NE cells, I performed loss-of-function and rescue experiments and found that the neuronal differentiation ability of NEs was adversely affected by knockdown of *BRN2* (Fig. 38). In either case, I observed only a few SMA-positive cells, and the majority of the cell population was neither βIII-tubulin nor SMA positive (Fig. 38). Overexpression of BRN2 in parallel with CHD7 knockdown rescued the aberrant phenotype, as evidenced by the significantly enhanced capacity of these cells for differentiation into βIII-tubulin-positive neurons and the complete absence of SMA-positive cells (Fig. 39). Overexpression of SOX21 in CHD7-knockdown NE cells resulted in a partial increase in the number of βIII-tubulin-positive neurons while moderately suppressing the generation of SMA-positive cells (Fig. 39).

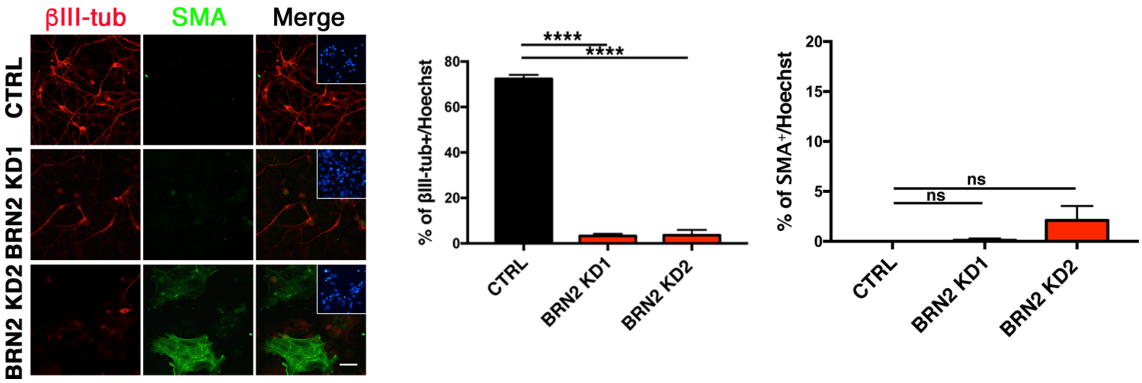


Figure 38: Neuronal differentiation ability was severely compromised in BRN2-knockdown cells. Immunostaining of control and BRN2-knockdown cells after neuronal differentiation (at day 22) for βIII-tubulin and SMA (left panel). Insets: Hoechst nuclear staining of each field. Scale bars, 50 µm. Quantification is shown in the right panel (n=3). Data are presented as the mean ± SEM (****p < 0.0001, two-tailed unpaired Student's t test).

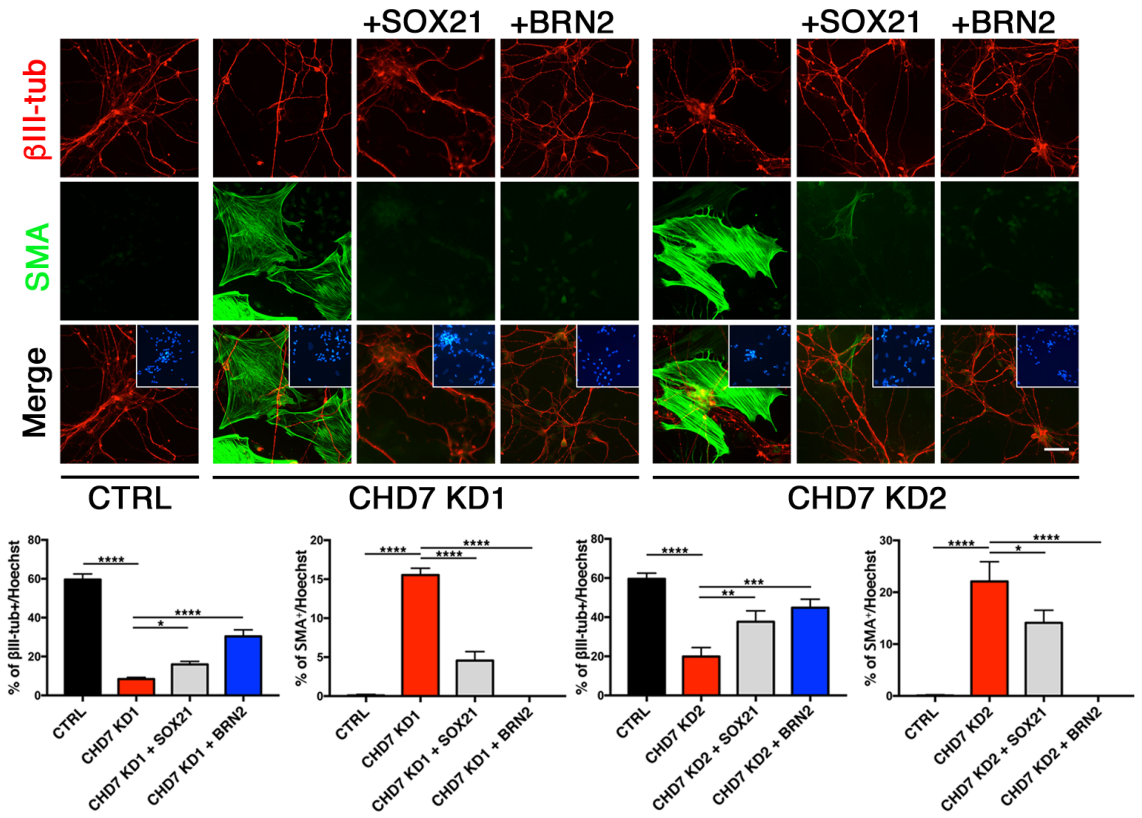


Figure 39: Overexpression of SOX21/BRN2 in CHD7-knockdown cells restored neural identity and neurogenic competence.
Immunostaining for βIII-tubulin (red) and SMA (green) in differentiated control, CHD7-knockdown, and SOX21 or BRN2 overexpressing CHD7-knockdown NE cells (left panel). Insets: Hoechst nuclear staining of each field. Quantification of the number of βIII-tubulin- and SMA-expressing cells under each condition is shown in the right panel (n=3). Quantification is presented as the mean ± SEM (*p < 0.05, **p < 0.01, ***p < 0.001 and ****p < 0.0001; two-tailed unpaired Student's t test).

I next asked whether overexpression of BRN2 in CHARGE-NE cells could restore neuronal differentiation capacity. As a result, I observed an increase in the population of βIII-tubulin-positive neurons (~30%) in BRN2-overexpressing CHARGE-NE cells compared with that in cells transduced with a control vector; however, these neurons appeared immature, as indicated by their relatively short processes (Fig. 40). Nevertheless, the expression of BRN2 inhibited the differentiation of SMA-positive cells (Fig. 40), suggesting a partial restoration of neural identity and neurogenic competence in CHARGE-NE cells. Although BRN2 exerts a stronger neurogenic effect than SOX21, our data further confirm the notion that SOX21 and BRN2 play a critical role in preserving the neurogenic potential of NE cells. Altogether, our findings indicate that, through modulation of super-enhancer activity, CHD7, SOX21 and BRN2 coordinately orchestrate NE-specific gene expression program during neuroepithelial-NC lineage commitment.

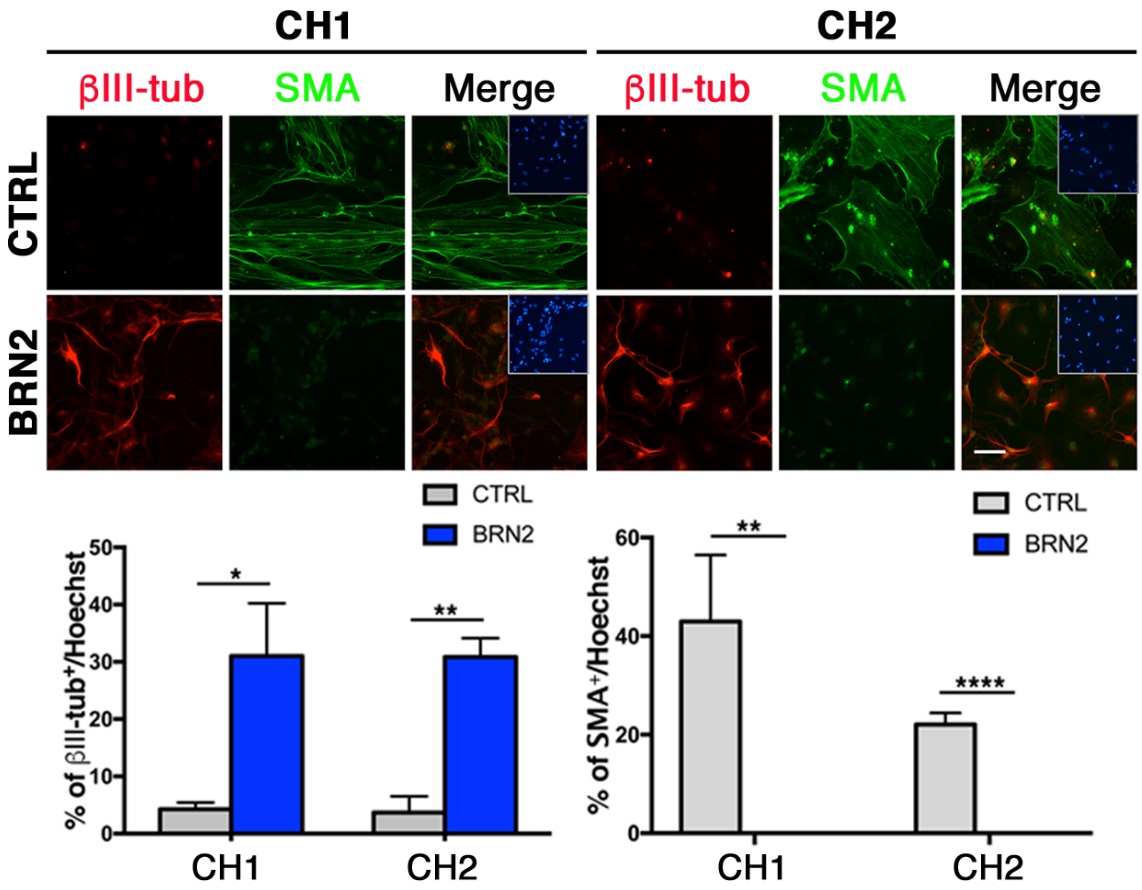


Figure 40: Overexpression of BRN2 in CHARGE-NE cells partially restored neural identity and neurogenic competence.

Immunostaining for βIII-tubulin (red) and SMA (green) in differentiated patient CHARGE-NE cells transduced with or without BRN2. Insets: Hoechst nuclear staining of each field. The number of βIII-tubulin- and SMA-expressing cells in each condition is shown in the right panel (n=3). Quantification is presented as the mean ± SEM (*$p < 0.05$, **$p < 0.01$ and ****$p < 0.0001$; one-way analysis of variance (ANOVA) with Bonferroni's post hoc test).

5. Discussion

In this study, I found that CHD7, the causative gene in CHARGE syndrome, is a master regulator of maintenance of the lineage-specific epigenome and cell type-specific gene expression that is integral to NE cell fate and CNS lineage commitment. An abundance of CHD7 is required to preserve the activation of cell type-specific enhancers for CNS lineage, maintaining the identity of CNS cell types irrespective of cellular state. Downregulation of CHD7 in NE cells caused a fate switch from NE to NC-like cells by a reduction in H3K27ac abundance at CNS-specific enhancers and conversely, an increase in H3K27ac abundance at non-CNS enhancers, which further instruct an abrupt change in the cell type-specific gene expression program. Finally, I showed that SOX21 and BRN2 were the downstream effectors of CHD7 function in controlling neurogenic competence. Most importantly, by using CHARGE syndrome patient-derived NE cells, our proposed mechanism has substantive implications in the pathogenesis of CHARGE syndrome.

Single-cell lineage analysis demonstrates that a single NE cell *in vivo* can differentiate into derivatives of either the CNS or NC (Bronner-Fraser and Fraser, 1989; Bronner-Fraser and Fraser, 1988; Brown and Storey, 2000; Frank and Sanes, 1991; Selleck and Bronner-Fraser, 1995). A recent study further demonstrated the conversion of mouse cortical neural stem cells at E14.5 to NCCs was possible by activating Sox9 while repressing Sox2 (Remboutsika et al., 2011). These studies indicate that the CNS is

not fully segregated from the NC even at mid-stage embryonic development; thus, the CNS must have mechanisms in place to maintain CNS lineage identity. Here, I provided three independent lines of evidence to support the notion that CHD7 maintains the current cell state and CNS lineage identity in NE cells. First, more than 50% of CHD7-knockdown NE cells lost the expression of NE identity markers, including SOX1 and DACH1 (Fig. 5). Second, morphological changes from rosette-like patterns to mesenchymal-like cell patterns were observed as early as day 5 after lentiviral-mediated CHD7 knockdown, and ectopic upregulation of NC master regulators SOX9 and SNAI2 was observed beginning at day 8 (Fig. 10). Furthermore, global transcriptome profiles of CHD7-knockdown NE cells revealed the manifestation of a NC-specific gene expression program (Fig. 9). Third, the neuronal differentiation rate decreased by more than 50% cells for CHD7-knockdown NE cells, which instead differentiated into smooth muscle cells (Fig. 6). In addition, I show that both *in vitro* and *in vivo*, NE cells express a higher level of CHD7 than NCCs (Fig. 2 and 3), suggesting that CHD7 may act as a segregation factor for the CNS and NC.

Mesenchymal stem cells (MSCs)/stromal cells share developmental similarities and differentiation potential with NCCs, consistent with the views that MSCs partially originate from NCCs (Fukuta et al., 2014; Nagoshi et al., 2008; Takashima et al., 2007). For instance, the craniofacial mesenchyme and a subpopulation of bone marrow MSCs developmentally originate from NCCs (Chai et al., 2000; Hagiwara et al., 2014; Isern et al., 2014; Wiszniak et al., 2015). Moreover, the mesoderm is a major source of the MSCs

that give rise to skeletal and connective tissues (Olsen et al., 2000). I found that the transcriptome profiles of CHD7-knockdown cells are equally consistent with MSCs. However, the current work was unable to distinguish the exact origin of these cells, specifically whether they are derived from the NC, neuroectoderm or mesoderm. In this context, a previous study showed that Sox1^{+} NE cells give rise to MSCs through a NC intermediate stage; however, this process accounts for merely 1.3% of the total MSC population (Takashima et al., 2007), suggesting that direct conversion from neuroepithelium to mesoderm is rather unlikely. Thus, further experiments are needed to determine the role of CHD7 in mesodermal lineage commitment.

Our study reveals a novel function for CHD7 in activating lineage-specific enhancers in NE cells. CHD7 preferentially binds to distal regulatory elements, particularly SEs, in human ESCs and mouse cerebellar granule neuron progenitors (Feng et al., 2017; Hnisz et al., 2013). Although SEs control cell identity, previous studies did not show disruption of CHD7 and presumably, binding of CHD7 to SEs, resulting in the loss of cell identity. Here, I present several lines of evidence that CHD7 regulates cell and lineage identities through maintenance of the active enhancer repertoire and a cell type-specific gene expression program. I observed a high level of CHD7 binding at SEs in which CHD7-associated genes are master regulators of neural development and neurogenesis (Fig. 24). Knockdown of CHD7 in NE cells correspondingly altered H3K27ac abundance surrounding SE regions (Fig. 28), consistent with the downregulation of 30% of SE-associated genes in CHD7-knockdown cells (Fig. 29) and

disablement of the NE-specific gene expression program (Fig. 9). I found that CHD7-bound distal regions are most enriched for CNS-specific enhancers (Fig. 25 and 26). In CHD7-knockdown cells, the average H3K27ac signal at CNS-specific enhancers but not non-CNS enhancers was decreased (Fig. 30 and 31). Our data show that CHD7 is directly responsible for enhancer activity, but our work is insufficient to address why NE cells undergo a fate switch to NC-like cells in the absence of CHD7, especially in light of our findings that non-CNS enhancers were overrepresented among CHD7-independent increases in H3K27ac regions upon CHD7 abrogation (Fig. 31).

I identified *SOX21* and *BRN2* as primary targets of CHD7 in NE cells (Fig. 32). At CHD7 peaks proximal to *SOX21* and *BRN2*, H3K27ac levels were reduced upon CHD7 knockdown (Fig. 33). The enhancer activities of these regions, as measured by a luciferase reporter assay, were markedly reduced in the presence of *CHD7* shRNAs (Fig. 34). Knockdown of BRN2 in NE cells adversely affected neuronal differentiation ability and milder defects with knockdown of SOX21, but no spontaneous smooth muscle differentiation was observed (Fig. 38). These results might be explained by the critical role that SOX21 and BRN2 play in conferring neurogenic competence to NE cells; however, they play a much less substantial role in blocking alternative lineages. Nevertheless, neuronal differentiation ability was restored, and smooth muscle differentiation was repressed in CHD7-knockdown cells overexpressing BRN2 or SOX21 (Fig. 39 and 40). The finding that the CHD7-driven regulatory program is essential for *in vitro* NE rosette cultures is of potential relevance in the development of 3D organoid-

based functional modeling of human brain development, given that the neuroepithelium is the fundamental building block of brain organoids.

Most importantly, our work shows that these novel functions of CHD7 in NE cells may underlie neurodevelopmental defects, particularly in the CNS of patients with CHARGE syndrome caused by frameshift or nonsense mutations of *CHD7*. I found that CHARGE-NE cells exhibited altered differentiation capacity similar to that of CHD7-knockdown cells (Fig. 14). By contrast, overexpression of $CHD7^{G1391fs}$ or $CHD7^{R1494X}$ in NE cells has no discernible phenotypes without any dominant negative effects (Fig. 17). Exogenous expression of CHD7 in CHARGE-NE cells enhanced neuronal differentiation capacity and inhibited smooth muscle cell differentiation (Fig. 15). Previous studies on lineage maintenance in the immune cells showed that the binding characteristic of a single TF to genomic regions relies on its nuclear concentration (Heinz et al., 2010). When a TF is expressed at lower levels, this TF relies strongly on its cognate partner TFs to bind to many genomic sites, thus establishing a different set of active enhancers than that established by a TF expressed at high levels. Consistently, this work shows that the level of CHD7 expression in CHARGE-NE cells is inadequate to maintain the CNS lineage-specific epigenome, ultimately leading to the loss of the NE-specific gene expression program (Fig. 16 and 18). This finding indicates an absolute requirement for the continuous presence of CHD7 in CNS development.

In conclusion, this study uncovers a previously unappreciated role for CHD7 in CNS lineage maintenance. By using iPSC-derived NE cell-based disease modeling, this

study sheds light on the molecular link between CNS anomalies and craniofacial malformations in CHARGE syndrome patients. These findings suggest that collapse of a cell type-specific gene expression program in the NE progenitor population that causes the ensuing lineage switch to multipotent NC-like cells is sufficient to affect a wide range of neural and NC derivatives throughout human fetal development. Thus, this study provides insights into the long-standing question regarding the causes of multiple anomalies in CHARGE syndrome patients. This study also highlights the feasibility of using human iPSC-based research to improve the understanding of human neurodevelopment and neurodevelopmental disorders, bringing us another step closer to decoding the complexity of the human CNS.

6. Acknowledgement

The completion of this endeavor signifies an important milestone in my life; not merely because I will be offered a doctoral degree, but having been trained in Japan and completed the training contributes to my greatest sense of accomplishment. The training has initiated a revolution within myself, which I believe has made me a better person/researcher than I was 8 years ago. These would not be possible without the passion and participation of so many people whose name may not all be enumerated here.

First of all, I would like to express my sincere appreciation and thanks to Professor Kinichi Nakashima, for nurturing me into a disciplined yet witty young scientist. His insistence on discipline, hard work, and passion for science is one of the pillars supporting my adventure in research.

I also would like to convey my deepest gratitude to Professor Yasumasa Bessho, who always goes the extra mile to help me. His constant encouragement, understanding, and coordination have greatly facilitated my completion. My sincere appreciation to Professor Masashi Kawaichi and Associate Professor Shoji Komai, the members of the advisory board, for their expert and valuable guidance extended to me.

In addition, I am also highly indebted to my current bosses, Professor Hideyuki Okano and Associate Professor Jun Kohyama, who offered me a second life in pursuing my doctorate. They have been conveying a spirit excitement in doing science consistently,

which inspired me to become fearless in science. Without their advice and support, this thesis would not have been possible.

At last but not the least, I would like to thank my wonderful husband, my beloved parents, my family members, my colleagues and my best friends for their unconditional love and patience in the last 8 years. They are always there for me no matter what.

I have no words to express my gratitude and appreciation, but my heart is still filled with the love and favors received from everyone.

7. References

Achilleos, A., and Trainor, P.A. (2012). Neural crest stem cells: discovery, properties and potential for therapy. Cell Res *22*, 288-304.

Adam, R.C., Yang, H., Rockowitz, S., Larsen, S.B., Nikolova, M., Oristian, D.S., Polak, L., Kadaja, M., Asare, A., Zheng, D.*, et al.* (2015). Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature *521*, 366-370.

Aramaki, M., Udaka, T., Kosaki, R., Makita, Y., Okamoto, N., Yoshihashi, H., Oki, H., Nanao, K., Moriyama, N., Oku, S.*, et al.* (2006). Phenotypic spectrum of CHARGE syndrome with CHD7 mutations. J Pediatr *148*, 410-414.

Bajpai, R., Chen, D.A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C.P., Zhao, Y., Swigut, T., and Wysocka, J. (2010). CHD7 cooperates with PBAF to control multipotent neural crest formation. Nature *463*, 958-962.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709-1712.

Bartholomew, B. (2014). Regulating the chromatin landscape: structural and mechanistic perspectives. Annu Rev Biochem *83*, 671-696.

Bergman, J.E., Janssen, N., Hoefsloot, L.H., Jongmans, M.C., Hofstra, R.M., and van Ravenswaaij-Arts, C.M. (2011). CHD7 mutations and CHARGE syndrome: the clinical implications of an expanding phenotype. J Med Genet *48*, 334-342.

Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci *31*, 35-40.

Boeger, H., Griesenbeck, J., and Kornberg, R.D. (2008). Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. Cell *133*, 716-726.

Bosman, E.A., Penn, A.C., Ambrose, J.C., Kettleborough, R., Stemple, D.L., and Steel, K.P. (2005). Multiple mutations in mouse Chd7 provide models for CHARGE syndrome. Hum Mol Genet *14*, 3463-3476.

Bowen, N.J., Fujita, N., Kajita, M., and Wade, P.A. (2004). Mi-2/NuRD:

multiple complexes for many purposes. Biochim Biophys Acta *1677*, 52-57.
Bronner-Fraser, M., and Fraser, S. (1989). Developmental potential of avian trunk neural crest cells in situ. Neuron *3*, 755-766.
Bronner-Fraser, M., and Fraser, S.E. (1988). Cell lineage analysis reveals multipotency of some avian neural crest cells. Nature *335*, 161-164.
Brown, J.M., and Storey, K.G. (2000). A region of the vertebrate neural plate in which neighbouring cells can adopt neural or epidermal fates. Curr Biol *10*, 869-872.
Bylund, M., Andersson, E., Novitch, B.G., and Muhr, J. (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. Nat Neurosci *6*, 1162-1168.
Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature *461*, 193-198.
Chai, Y., Jiang, X., Ito, Y., Bringas, P., Jr., Han, J., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. Development *127*, 1671-1679.
Cheung, M., and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. Development *130*, 5681-5693.
Cheung, M., Chaboissier, M.C., Mynett, A., Hirst, E., Schedl, A., and Briscoe, J. (2005). The transcriptional control of trunk neural crest induction, survival, and delamination. Dev Cell *8*, 179-192.
Cohen-Carmon, D., and Meshorer, E. (2012). Polyglutamine (polyQ) disorders: the chromatin connection. Nucleus *3*, 433-441.
Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.
Corsten-Janssen, N., Saitta, S.C., Hoefsloot, L.H., McDonald-McGinn, D.M., Driscoll, D.A., Derks, R., Dickinson, K.A., Kerstjens-Frederikse, W.S., Emanuel, B.S., Zackai, E.H.*, et al.* (2013). More Clinical Overlap between 22q11.2 Deletion Syndrome and CHARGE Syndrome than Often Anticipated. Mol Syndromol *4*, 235-245.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.*

(2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A *107*, 21931-21936.

Curchoe, C.L., Russo, J., and Terskikh, A.V. (2012). hESC derived neuro-epithelial rosettes recapitulate early mammalian neurulation events; an in vitro model. Stem Cell Res *8*, 239-246.

Dawson, M.A., and Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. Cell *150*, 12-27.

Dechassa, M.L., Sabri, A., Pondugula, S., Kassabov, S.R., Chatterjee, N., Kladde, M.P., and Bartholomew, B. (2010). SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. Mol Cell *38*, 590-602.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602-607.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Engelen, E., Akinci, U., Bryne, J.C., Hou, J., Gontan, C., Moen, M., Szumska, D., Kockx, C., van Ijcken, W., Dekkers, D.H.*, et al.* (2011). Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. Nat Genet *43*, 607-611.

Falk, A., Koch, P., Kesavan, J., Takashima, Y., Ladewig, J., Alexander, M., Wiskow, O., Tailor, J., Trotter, M., Pollard, S.*, et al.* (2012). Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. PLoS One *7*, e29597.

Feng, W., Kawauchi, D., Korkel-Qu, H., Deng, H., Serger, E., Sieber, L., Lieberman, J.A., Jimeno-Gonzalez, S., Lambo, S., Hanna, B.S.*, et al.* (2017). Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. Nat Commun *8*, 14758.

Feng, W., Khan, M.A., Bellvis, P., Zhu, Z., Bernhardt, O., Herold-Mende, C., and Liu, H.K. (2013). The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors. Cell Stem Cell *13*,

62-72.
Frank, E., and Sanes, J.R. (1991). Lineage of neurons and glia in chick dorsal root ganglia: analysis in vivo with a recombinant retrovirus. Development *111*, 895-908.
Fukuta, M., Nakai, Y., Kirino, K., Nakagawa, M., Sekiguchi, K., Nagata, S., Matsumoto, Y., Yamamoto, T., Umeda, K., Heike, T.*, et al.* (2014). Derivation of mesenchymal stromal cells from pluripotent stem cells through a neural crest lineage using small molecule compounds with defined media. PLoS One *9*, e112291.
Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E.*, et al.* (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863-868.
Gibbons, R.J., McDowell, T.L., Raman, S., O'Rourke, D.M., Garrick, D., Ayyub, H., and Higgs, D.R. (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet *24*, 368-371.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A.*, et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442-451.
Gomez-Del Arco, P., Perdiguero, E., Yunes-Leites, P.S., Acin-Perez, R., Zeini, M., Garcia-Gomez, A., Sreenivasan, K., Jimenez-Alcazar, M., Segales, J., Lopez-Maderuelo, D.*, et al.* (2016). The Chromatin Remodeling Complex Chd4/NuRD Controls Striated Muscle Identity and Metabolic Homeostasis. Cell Metab *23*, 881-892.
Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J.J., Hession, C., Zhang, F., and Regev, A. (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. Science *353*, 925-928.
Hagiwara, K., Obayashi, T., Sakayori, N., Yamanishi, E., Hayashi, R., Osumi, N., Nakazawa, T., and Nishida, K. (2014). Molecular and cellular features of murine craniofacial and trunk neural crest cells as stem cell-like cells. PLoS

One *9*, e84072.
Hall, J.A., and Georgel, P.T. (2007). CHD proteins: a diverse family with strong ties. Biochem Cell Biol *85*, 463-476.
He, D., Marie, C., Zhao, C., Kim, B., Wang, J., Deng, Y., Clavairoly, A., Frah, M., Wang, H., He, X.*, et al.* (2016). Chd7 cooperates with Sox10 and regulates the onset of CNS myelination and remyelination. Nat Neurosci *19*, 678-689.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
Higuchi, R., Krummel, B., and Saiki, R.K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res *16*, 7351-7367.
Hill, R.E., Jones, P.F., Rees, A.R., Sime, C.M., Justice, M.J., Copeland, N.G., Jenkins, N.A., Graham, E., and Davidson, D.R. (1989). A new family of mouse homeo box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox-7.1. Genes Dev *3*, 26-37.
Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.
Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.
Hurd, E.A., Poucher, H.K., Cheng, K., Raphael, Y., and Martin, D.M. (2010). The ATP-dependent chromatin remodeling enzyme CHD7 regulates pro-neural gene expression and neurogenesis in the inner ear. Development *137*, 3139-3150.
Isern, J., Garcia-Garcia, A., Martin, A.M., Arranz, L., Martin-Perez, D., Torroja, C., Sanchez-Cabo, F., and Mendez-Ferrer, S. (2014). The neural crest is a source of mesenchymal stem cells with specialized hematopoietic stem cell niche function. Elife *3*, e03696.
Isoda, M., Kohyama, J., Iwanami, A., Sanosaka, T., Sugai, K., Yamaguchi, R., Matsumoto, T., Nakamura, M., and Okano, H. (2016a). Robust production of human neural cells by establishing neuroepithelial-like stem cells from

peripheral blood mononuclear cell-derived feeder-free iPSCs under xeno-free conditions. Neurosci Res.

Isoda, M., Kohyama, J., Iwanami, A., Sanosaka, T., Sugai, K., Yamaguchi, R., Matsumoto, T., Nakamura, M., and Okano, H. (2016b). Robust production of human neural cells by establishing neuroepithelial-like stem cells from peripheral blood mononuclear cell-derived feeder-free iPSCs under xeno-free conditions. Neurosci Res *110*, 18-28.

Iyer, V., and Struhl, K. (1995). Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. EMBO J *14*, 2570-2579.

Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol *43*, 1565-1575.

Janssen, N., Bergman, J.E., Swertz, M.A., Tranebjaerg, L., Lodahl, M., Schoots, J., Hofstra, R.M., van Ravenswaaij-Arts, C.M., and Hoefsloot, L.H. (2012). Mutation update on the CHD7 gene involved in CHARGE syndrome. Hum Mutat *33*, 1149-1160.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet *10*, 161-172.

Jiang, X., Zhou, Y., Xian, L., Chen, W., Wu, H., and Gao, X. (2012). The mutation in Chd7 causes misexpression of Bmp4 and developmental defects in telencephalic midline. Am J Pathol *181*, 626-641.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497-1502.

Jones, K.M., Saric, N., Russell, J.P., Andoniadou, C.L., Scambler, P.J., and Basson, M.A. (2015). CHD7 maintains neural stem cell quiescence and prevents premature stem cell depletion in the adult hippocampus. Stem Cells *33*, 196-210.

Kamakaka, R.T., and Biggins, S. (2005). Histone variants: deviants? Genes Dev *19*, 295-310.

Keating, S.T., and El-Osta, A. (2012). Chromatin modifications associated with diabetes. J Cardiovasc Transl Res *5*, 399-412.
Kim, H., and Kim, J.S. (2014). A guide to genome engineering with programmable nucleases. Nat Rev Genet *15*, 321-334.
Koch, P., Opitz, T., Steinbeck, J.A., Ladewig, J., and Brustle, O. (2009). A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. Proc Natl Acad Sci U S A *106*, 3225-3230.
Komoike, Y., Matsuoka, M., and Kosaki, K. (2013). Potential teratogenicity of methimazole: exposure of zebrafish embryos to methimazole causes similar developmental anomalies to human methimazole embryopathy. Birth Defects Res B Dev Reprod Toxicol *98*, 222-229.
Kornberg, R.D., and Klug, A. (1981). The nucleosome. Sci Am *244*, 52-64.
Lalani, S.R., Safiullah, A.M., Molinari, L.M., Fernbach, S.D., Martin, D.M., and Belmont, J.W. (2004). SEMA3E mutation in a patient with CHARGE syndrome. J Med Genet *41*, e94.
Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373-379.
Layman, W.S., McEwen, D.P., Beyer, L.A., Lalani, S.R., Fernbach, S.D., Oh, E., Swaroop, A., Hegg, C.C., Raphael, Y., Martens, J.R.*, et al.* (2009). Defects in neural stem cell proliferation and olfaction in Chd7 deficient mice indicate a mechanism for hyposmia in human CHARGE syndrome. Hum Mol Genet *18*, 1909-1923.
Li, W., Sun, W., Zhang, Y., Wei, W., Ambasudhan, R., Xia, P., Talantova, M., Lin, T., Kim, J., Wang, X.*, et al.* (2011). Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. Proc Natl Acad Sci U S A *108*, 8299-8304.
Liu, X., Lee, C.K., Granek, J.A., Clarke, N.D., and Lieb, J.D. (2006). Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. Genome Res *16*, 1517-1528.

Loven, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320-334.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luk, E., Ranjan, A., Fitzgerald, P.C., Mizuguchi, G., Huang, Y., Wei, D., and Wu, C. (2010). Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell *143*, 725-736.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol *31*, 833-838.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol *60*, 174-182.

Moustafa-Hawash, N., Smolkin, T., Ilivitzki, A., Zimberg-Bossira, A., Gildish, A., Gershoni-Baruch, R., and Makhoul, I.R. (2012). CHARGE syndrome with del(3)(p13p21): expanding the genotype. Isr Med Assoc J *14*, 133-134.

Nagoshi, N., Shibata, S., Kubota, Y., Nakamura, M., Nagai, Y., Satoh, E., Morikawa, S., Okada, Y., Mabuchi, Y., Katoh, H.*, et al.* (2008). Ontogeny and multipotency of neural crest-derived stem cells in mouse bone marrow, dorsal root ganglia, and whisker pad. Cell Stem Cell *2*, 392-403.

Nakagawa, M., Taniguchi, Y., Senda, S., Takizawa, N., Ichisaka, T., Asano, K., Morizane, A., Doi, D., Takahashi, J., Nishizawa, M.*, et al.* (2014). A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. Sci Rep *4*, 3594.

Nakashima, K., Yanagisawa, M., Arakawa, H., Kimura, N., Hisatsune, T., Kawabata, M., Miyazono, K., and Taga, T. (1999). Synergistic signaling in fetal brain by STAT3-Smad1 complex bridged by p300. Science *284*, 479-482.
Nieto, M.A., Sargent, M.G., Wilkinson, D.G., and Cooke, J. (1994). Control of cell behavior during vertebrate development by Slug, a zinc finger gene. Science *264*, 835-839.
Nikitina, N., Sauka-Spengler, T., and Bronner-Fraser, M. (2008). Dissecting early regulatory relationships in the lamprey neural crest gene network. Proc Natl Acad Sci U S A *105*, 20083-20088.
Okita, K., Yamakawa, T., Matsumura, Y., Sato, Y., Amano, N., Watanabe, A., Goshima, N., and Yamanaka, S. (2013). An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. Stem Cells *31*, 458-466.
Okuno, H., Renault Mihara, F., Ohta, S., Fukuda, K., Kurosawa, K., Akamatsu, W., Sanosaka, T., Kohyama, J., Hayashi, K., Nakajima, K.*, et al.* (2017). CHARGE syndrome modeling using patient-iPSCs reveals defective migration of neural crest cells harboring CHD7 mutations. Elife *6*.
Olsen, B.R., Reginato, A.M., and Wang, W. (2000). Bone development. Annu Rev Cell Dev Biol *16*, 191-220.
Pagon, R.A., Graham, J.M., Jr., Zonana, J., and Yong, S.L. (1981). Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. J Pediatr *99*, 223-227.
Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell *144*, 200-213.
Petty, E., and Pillus, L. (2013). Balancing chromatin remodeling and histone modifications in transcription. Trends Genet *29*, 621-629.
Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.
Remboutsika, E., Elkouris, M., Iulianella, A., Andoniadou, C.L., Poulou, M., Mitsiadis, T.A., Trainor, P.A., and Lovell-Badge, R. (2011). Flexibility of neural stem cells. Front Physiol *2*, 16.

Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: unexpected roles for chromatin. Nat Rev Genet *14*, 347-359.
Sandberg, M., Kallstrom, M., and Muhr, J. (2005). Sox21 promotes the progression of vertebrate neurogenesis. Nat Neurosci *8*, 995-1001.
Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol *32*, 347-355.
Sanlaville, D., Etchevers, H.C., Gonzales, M., Martinovic, J., Clement-Ziza, M., Delezoide, A.L., Aubry, M.C., Pelet, A., Chemouny, S., Cruaud, C.*, et al.* (2006). Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. J Med Genet *43*, 211-217.
Sanlaville, D., and Verloes, A. (2007). CHARGE syndrome: an update. Eur J Hum Genet *15*, 389-399.
Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. Nat Rev Mol Cell Biol *6*, 139-149.
Sauka-Spengler, T., and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. Nat Rev Mol Cell Biol *9*, 557-568.
Savic, D., Partridge, E.C., Newberry, K.M., Smith, S.B., Meadows, S.K., Roberts, B.S., Mackiewicz, M., Mendenhall, E.M., and Myers, R.M. (2015). CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res *25*, 1581-1589.
Scherson, T., Serbedzija, G., Fraser, S., and Bronner-Fraser, M. (1993). Regulative capacity of the cranial neural tube to form neural crest. Development *118*, 1049-1062.
Schnetz, M.P., Bartels, C.F., Shastri, K., Balasubramanian, D., Zentner, G.E., Balaji, R., Zhang, X., Song, L., Wang, Z., Laframboise, T.*, et al.* (2009). Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. Genome Res *19*, 590-601.
Schnetz, M.P., Handoko, L., Akhtar-Zaidi, B., Bartels, C.F., Pereira, C.F., Fisher, A.G., Adams, D.J., Flicek, P., Crawford, G.E., Laframboise, T.*, et al.* (2010). CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. PLoS Genet *6*, e1001023.

Schulz, Y., Wehner, P., Opitz, L., Salinas-Riester, G., Bongers, E.M., van Ravenswaaij-Arts, C.M., Wincent, J., Schoumans, J., Kohlhase, J., Borchers, A., *et al.* (2014). CHD7, the gene mutated in CHARGE syndrome, regulates genes involved in neural crest cell guidance. Hum Genet *133*, 997-1009.

Selleck, M.A., and Bronner-Fraser, M. (1995). Origins of the avian neural crest: the role of neural plate-epidermal interactions. Development *121*, 525-538.

Shimojo, D., Onodera, K., Doi-Torii, Y., Ishihara, Y., Hattori, C., Miwa, Y., Tanaka, S., Okada, R., Ohyama, M., Shoji, M., *et al.* (2015). Rapid, efficient, and simple motor neuron differentiation from human pluripotent stem cells. Mol Brain *8*, 79.

Siebert, J.R., Graham, J.M., Jr., and MacDonald, C. (1985). Pathologic features of the CHARGE association: support for involvement of the neural crest. Teratology *31*, 331-336.

Simoes-Costa, M., and Bronner, M.E. (2013). Insights into neural crest development and evolution from genomic analysis. Genome Res *23*, 1069-1080.

Sims, J.K., and Wade, P.A. (2011). SnapShot: Chromatin remodeling: CHD. Cell *144*, 626-626 e621.

Snijders Blok, C., Corsten-Janssen, N., FitzPatrick, D.R., Romano, C., Fichera, M., Vitello, G.A., Willemsen, M.H., Schoots, J., Pfundt, R., van Ravenswaaij-Arts, C.M., *et al.* (2014). Definition of 5q11.2 microdeletion syndrome reveals overlap with CHARGE syndrome and 22q11 deletion syndrome phenotypes. Am J Med Genet A *164A*, 2843-2848.

Sperry, E.D., Hurd, E.A., Durham, M.A., Reamer, E.N., Stein, A.B., and Martin, D.M. (2014). The chromatin remodeling protein CHD7, mutated in CHARGE syndrome, is necessary for proper craniofacial and tracheal development. Dev Dyn *243*, 1055-1066.

Stokes, D.G., and Perry, R.P. (1995). DNA-binding and chromatin localization properties of CHD1. Mol Cell Biol *15*, 2745-2753.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat Struct Mol Biol *20*, 267-273.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K.,

and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takashima, Y., Era, T., Nakao, K., Kondo, S., Kasuga, M., Smith, A.G., and Nishikawa, S. (2007). Neuroepithelial cells supply an initial transient wave of MSC differentiation. Cell *129*, 1377-1388.

Thomas, S., Thomas, M., Wincker, P., Babarit, C., Xu, P., Speer, M.C., Munnich, A., Lyonnet, S., Vekemans, M., and Etchevers, H.C. (2008). Human neural crest cells display molecular and phenotypic hallmarks of stem cells. Hum Mol Genet *17*, 3411-3425.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-1041.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res *35*, D88-92.

Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R., Hurst, J.A., de Vries, B.B., Janssen, I.M., van der Vliet, W.A., Huys, E.H., de Jong, P.J., Hamel, B.C.*, et al.* (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet *36*, 955-957.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.

Williamson, K.A., Hever, A.M., Rainger, J., Rogers, R.C., Magee, A., Fiedler, Z., Keng, W.T., Sharkey, F.H., McGill, N., Hill, C.J.*, et al.* (2006). Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. Hum Mol Genet *15*, 1413-1422.

Wiszniak, S., Mackenzie, F.E., Anderson, P., Kabbara, S., Ruhrberg, C., and Schwarz, Q. (2015). Neural crest cell-derived VEGF promotes embryonic jaw extension. Proc Natl Acad Sci U S A *112*, 6086-6091.

Xue, Y., Wong, J., Moreno, G.T., Young, M.K., Cote, J., and Wang, W. (1998). NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell *2*, 851-861.

Yu, M., Yang, W., Ni, T., Tang, Z., Nakadai, T., Zhu, J., and Roeder, R.G. (2015). RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. Science *350*, 1383-1386.

Yu, T., Meiners, L.C., Danielsen, K., Wong, M.T., Bowler, T., Reinberg, D., Scambler, P.J., van Ravenswaaij-Arts, C.M., and Basson, M.A. (2013). Deregulated FGF and homeotic gene expression underlies cerebellar vermis hypoplasia in CHARGE syndrome. Elife *2*, e01305.

Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A *108*, 1531-1536.

Zentner, G.E., Layman, W.S., Martin, D.M., and Scacheri, P.C. (2010). Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. Am J Med Genet A *152A*, 674-686.

Zhou, Q., Li, T., and Price, D.H. (2012). RNA polymerase II elongation control. Annu Rev Biochem *81*, 119-143.