



OPEN

# Chromatin remodeler CHD7 targets active enhancer region to regulate cell type-specific gene expression in human neural crest cells

Tsukasa Sanosaka, Hironobu Okuno, Noriko Mizota, Tomoko Andoh-Noda, Miki Sato, Ryo Tomooka, Satoe Banno, Jun Kohyama & Hideyuki Okano

A mutation in the chromatin remodeler chromodomain helicase DNA-binding 7 (CHD7) gene causes the multiple congenital anomaly CHARGE syndrome. The craniofacial anomalies observed in CHARGE syndrome are caused by dysfunctions of neural crest cells (NCCs), which originate from the neural tube. However, the mechanism by which CHD7 regulates the function of human NCCs (hNCCs) remains unclear. We aimed to characterize the cis-regulatory elements governed by CHD7 in hNCCs by analyzing genome-wide ChIP-Seq data and identifying hNCC-specific CHD7-binding profiles. We compared CHD7-binding regions among cell types, including human induced pluripotent stem cells and human neuroepithelial cells, to determine the comprehensive properties of CHD7-binding in hNCCs. Importantly, analysis of the hNCC-specific CHD7-bound region revealed transcription factor AP-2 $\alpha$  as a potential co-factor facilitating the cell type-specific transcriptional program in hNCCs. CHD7 was strongly associated with active enhancer regions, permitting the expression of hNCC-specific genes to sustain the function of hNCCs. Our findings reveal the regulatory mechanisms of CHD7 in hNCCs, thus providing additional information regarding the transcriptional programs in hNCCs.

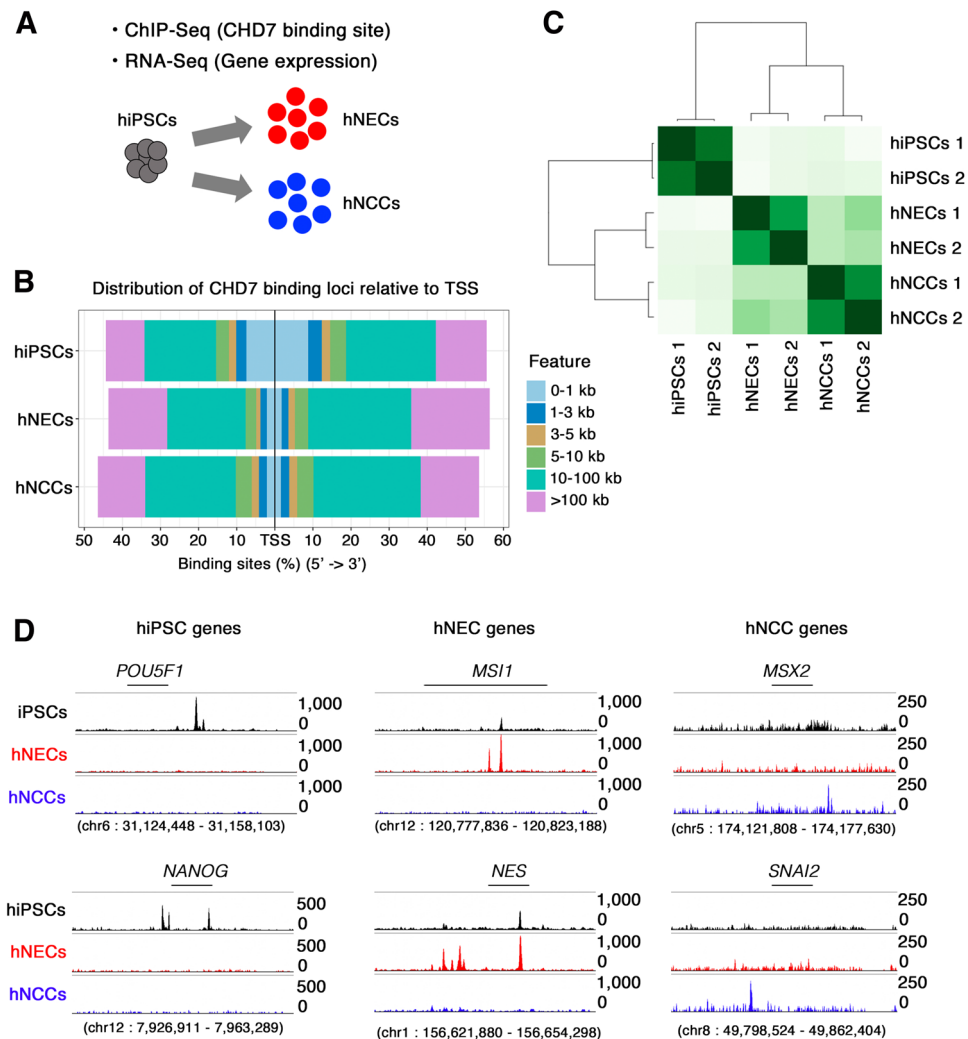
Cell type-specific transcriptional responses regulate cellular identity, which is dynamically coordinated during the developmental process<sup>1</sup>. Enhancers are genetic elements that were historically defined as DNA sequences that potentiate the cell type-specific transcription of transcriptional start sites (TSSs) located distally to the enhancer regions<sup>2</sup>. Although enhancers are responsible for exerting cell type-specific gene expression, they need to be accessible by adapting the local chromatin structure and enabling a precise response to external stimuli during development<sup>3,4</sup>. Accordingly, chromatin-based epigenetic alterations are closely linked to enhancer activity. The importance of chromatin regulation is also evidenced by the fact that defects in chromatin regulators cause genetic developmental disorders called chromatinopathies<sup>5,6</sup>. Notably, a genetic mutation in chromatin remodelers was observed in a subset of chromatinopathies, including ATRX syndrome, Floating-Harbor syndrome, and CHARGE syndrome<sup>5</sup>.

CHARGE syndrome is a congenital disorder caused by a heterozygous mutation in *CHD7*<sup>7,8</sup>. 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<sup>9</sup>.

We previously demonstrated an epigenetic landscape regulated by CHD7 in human induced pluripotent stem cell (hiPSC)-derived neuroepithelial cells (hNECs)<sup>10</sup>. Notably, CHD7 binds and regulates the activity of central nervous system-specific enhancers to control the differentiation capacity of hNECs<sup>10</sup>, indicating an essential role of CHD7 in regulating enhancers, at least in the central nervous system.

Dysfunction of neural crest cells (NCCs) is a primary cause of CHARGE syndrome<sup>8</sup>, and the role of CHD7 in regulating the epigenetic signature of NCCs has thus been of interest. The pioneering work by Wysocka's team using an in vitro differentiation system from human embryonic stem cells demonstrated that the formation of NCCs required the function of CHD7<sup>11</sup>. Additionally, we previously demonstrated defective migration of human NCCs (hNCCs) derived from iPSCs from patients with CHARGE syndrome<sup>12</sup>. These results provide potential explanations for the pathology of CHARGE syndrome. However, although CHD7 has been shown to

Department of Physiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. email: jkohyama@a7.keio.jp; hidokano@keio.jp



**Figure 1.** Identification of cell type-specific CHD7 targets. **(A)** Schematic illustration of the current study. **(B)** Distribution of CHD7-binding loci from TSS regions in hiPSCs, hNECs, and hNCCs. **(C)** Hierarchical clustering using Pearson's correlation analysis across the entire sample set. **(D)** ChIP-seq signals for CHD7 at genomic loci specific to hiPSCs (*POU5F1* and *NANOG*) (left panel), hNECs (*MSI1* and *NES*) (middle panel), and hNCCs (*MSX2* and *SNAI2*) (right panel) in each cell line.

be essential for the function and development of hNCCs, the comprehensive genomic map governed by CHD7 in hNCCs is still unclear.

We previously optimized the experimental conditions of chromatin immunoprecipitation (ChIP) followed by massive parallel DNA sequencing (ChIP-seq) for CHD7 and successfully generated CHD7-binding profiles in hNECs<sup>10</sup>. However, detailed information on the cell type-specific role of CHD7, especially in hNCCs, was not determined. In the current study, we aimed to compare CHD7-binding profiles in hiPSCs, hNECs, and hNCCs to identify hNCC-specific targets of CHD7. These findings highlight the fundamental role of CHD7 in cell type-specific gene regulation in hNCCs.

## Results

**Identification of cell type-specific CHD7 targets by ChIP-seq in hiPSCs, hNECs, and hNCCs.** The pivotal function of CHD7 in the development and function of various tissues depends on the cellular context<sup>10,11,13–18</sup>, and identifying the cell-specific targets of CHD7 is thus necessary to elucidate its function. We therefore investigated the function of CHD7 in hNCCs using the ChIP-seq profiles of CHD7 in hiPSCs, hNECs, and hNCCs<sup>10</sup> with transcriptome information using published datasets (Fig. 1A and Supplementary Table S1). We initially identified the CHD7-binding profiles using the Model-based Analysis of the ChIP-Seq (MACS2) algorithm and examined the distribution of CHD7-binding peaks in the genome (Fig. 1B and Supplementary Fig. S1). The average distribution of CHD7 peaks displayed similar genomic distributions, peaking at the TSS and transcription end site (TES) in all the tested cells in the current analysis (Supplementary Fig. S1). Most CHD7-binding peaks were located distally from the TSS in hNECs and hNCCs (Fig. 1B), but were preferentially located in the proximal region of the TSS in hiPSCs. These results indicated that the involvement of

CHD7 in gene regulation is cell-type dependent (Fig. 1B). We further selected cell type-specific CHD7-binding peaks using DiffBind<sup>19</sup>. Hierarchical clustering analysis of CHD7-binding regions revealed differential CHD7 occupancy among hiPSCs, hNECs, and hNCCs (Fig. 1C). We also found clear separation of cell type-specific CHD7 peaks in representative gene loci (Fig. 1D).

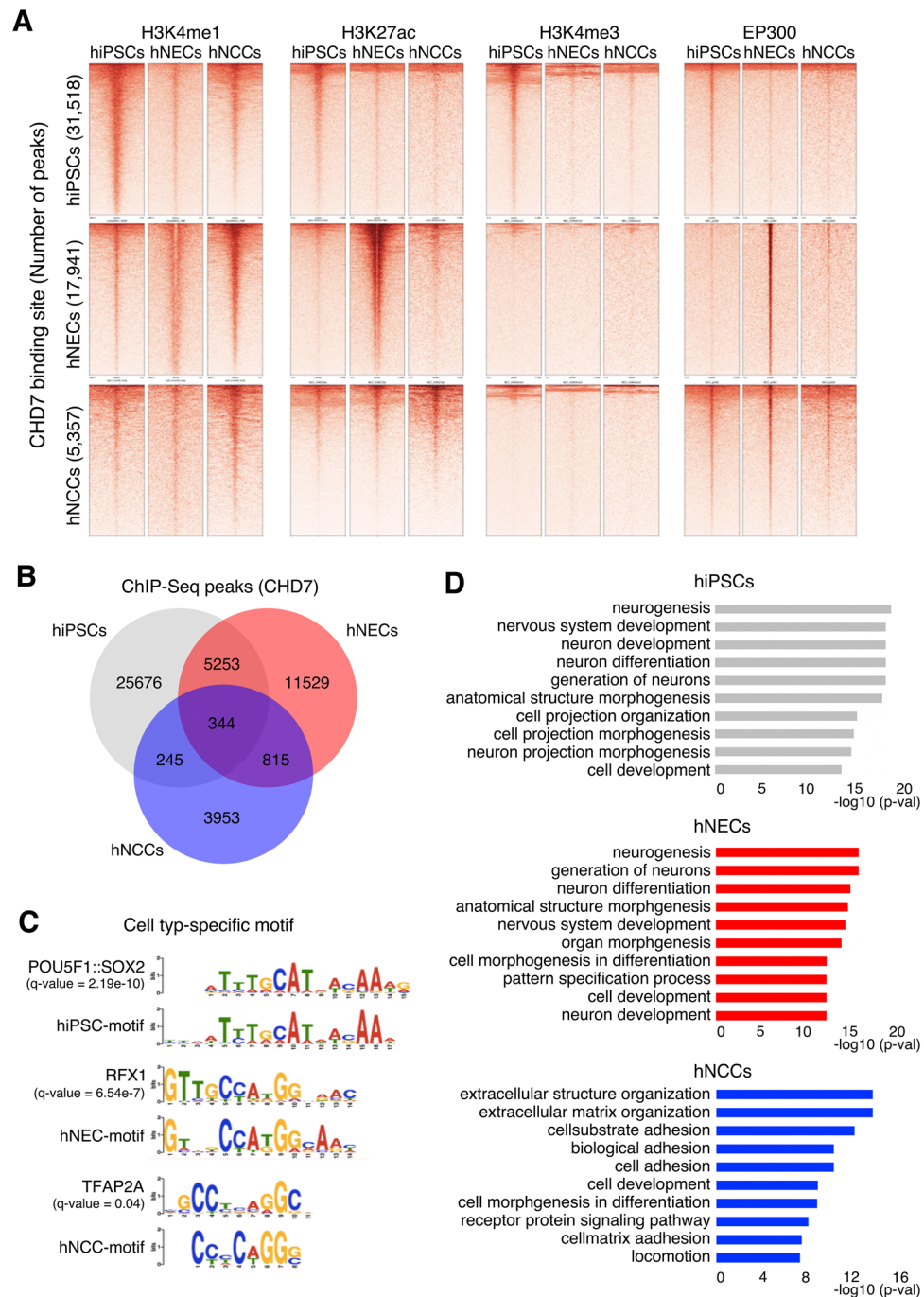
**CHD7 preferentially bound to active enhancers in hNCCs.** To gain further insights into the functional properties of CHD7, we cross-referenced CHD7-bound regions with the genome-wide chromatin state of histone modification in hiPSCs, hNECs, and hNCCs (Fig. 2A). We generated chromatin landscapes of cell type-specific CHD7-bound regions, including H3K4me1, H3K27ac, H3K4me3, and EP300. Bioinformatics analyses revealed a striking overlap between CHD7-bound regions and p300, H3K27ac, and H3K4me1, as markers indicative of active enhancers in hNECs and hNCCs (Fig. 2A). In contrast, CHD7 binding in hiPSCs was most likely to occur within the promoter region, evidenced by the correlation with H3K4me3, as described previously<sup>20,21</sup>. We aimed to identify the cell type-specific function of CHD7 and we therefore compared the locations of ChIP-seq peaks among hiPSCs, hNECs, and hNCCs. Although there were some shared peaks, most peaks were specific for each cell type (25,676 (81.5%) of 31,518 peaks in hiPSCs, 11,529 (64.3%) of 17,941 peaks in hNECs, and 3,953 (73.8%) of 5,357 peaks in hNCCs) (Fig. 2B and Supplementary Fig. S1). A total of 1,159 peaks were shared between hNECs and hNCCs, and 16,782 and 4,198 peaks were specific to hNECs and hNCCs, respectively, suggesting that CHD7 might play cell type-specific roles (Fig. 2B). CHD7 functions as a chromatin remodeler and requires interaction with DNA-bound transcription factors<sup>9</sup>. We thus speculated that the detected binding sites might aid the detection of such transcription factors in each cell type. We therefore performed motif analysis using MEME software to define cell type-specific peaks in each cell type<sup>22,23</sup> (Supplementary Fig. S1). The known binding motifs that matched the de novo motif in the cell type-specific peaks were target sequences for POU5F1:SOX2<sup>24</sup> in hiPSCs, regulatory factor X1<sup>25</sup> in hNECs, and transcription factor AP-2α (TFAP2A)<sup>26</sup> in hNCCs, respectively (Fig. 2C); notably, each transcription factor is a known regulator of each cell type. For hNCCs, TFAP2A is required for both neural plate-border induction and specification of NCCs<sup>26</sup>. TFAP2A also acts as a pioneer transcription factor that engages silent genes, helping to initiate neural crest (NC) specification<sup>27</sup>. As noted above, dysfunction of hNCCs has been reported to be involved in the pathogenesis of CHARGE syndrome<sup>12,28,29</sup>. A potential link between CHD7 and TFAP2A might thus provide potential mechanistic insights into the pathogenesis of CHARGE-derived hNCCs.

Gene Ontology (GO) analysis for hiPSCs and the hNEC-specific CHD7-regulated region revealed a similar trend, indicating that the sites were associated with neural differentiation, including “neurogenesis,” “nervous system development,” and “neuron differentiation” (Fig. 2D). In contrast, hNCC-specific CHD7-regulated regions were associated with mesenchymal properties of the NC, including “extracellular structure organization,” “extracellular matrix organization,” and “biological adhesion,” indicating successful identification of cell type-specific CHD7-bound regions in the current analyses (Fig. 2D).

**CHD7 binding predicted cell type-specific gene expression in hNCCs.** A previous study showed that super-enhancers (SEs) conferred cellular identities by regulating tissue-specific gene expression<sup>1,30,31</sup>. Moreover, we previously identified CHD7 as one of the components enriched in SEs across the hNEC genome<sup>10</sup>. This prompted us to determine if CHD7 was also enriched in SEs in hNCCs. We delineated SEs in hNCCs based on H3K27ac abundance using the ROSE algorithm, and identified 1,394 SEs (Fig. 3A). The identified SEs were located proximal to genes encoding key regulators of hNCC development, including SOX9, MSX1, and MSX2<sup>32–34</sup> (Fig. 3A, Supplementary Table S2). Furthermore, we found that CHD7 bound to SEs in hNCCs (Fig. 3B). To further evaluate the enhancer activity of CHD7-bound regions, we cross-referenced these regions with the VISTA Enhancer Browser<sup>35</sup>, a database of tissue-specific human enhancers with validated activity in mice. For example, CHD7 binding at SE-associated genes in hNCCs, including *MSX2*<sup>36</sup> and *GPC1*<sup>37</sup> loci, coincided with the presence of annotated enhancers that exhibited intense activity in NCC-associated tissues (Fig. 3C). To examine enhancer activity of these two loci in hNCCs, we generated luciferase reporter system containing the predicted enhancer region of *MSX2* or *GPC1* (Fig. 3D). hiPSCs, hNECs, and hNCCs were transfected with the luciferase reporter. As a reference, these cells were transfected with an empty vector. Then, the enhancer activity was evaluated by the degree of increased reporter activity with respect to the empty vector. As expected, higher reporter activity was detected in hNCCs compared to hiPSC and hNECs (Fig. 3D). In addition, the VISTA Enhancer Browser showed that CHD7-binding sites were most enriched for regions that displayed enhancer activity in NC derivatives, including dorsal root ganglia, heart, and ear (Fig. 3E). The identification of hNCC-specific CHD7-enriched enhancers suggests that CHD7 regulates hNCC-specificity of its corresponding genes.

We further elucidated the role of CHD7 in transcriptional regulation by examining the expression of cell type-specific CHD7 target genes. First, we determined cell type-specific CHD7 target genes based on the ChIP-seq information for hiPSCs, hNECs, and hNCCs (Supplementary Table S3). We then examined the expression of hNCC-specific CHD7 target genes in each cell type. As expected, the expression levels of these genes were higher in hNCCs (Fig. 4A). We also performed a correlation analysis and showed that the expression of cell type-specific CHD7 target genes was positively correlated with the expression of cell type-specific genes (Fig. 4B).

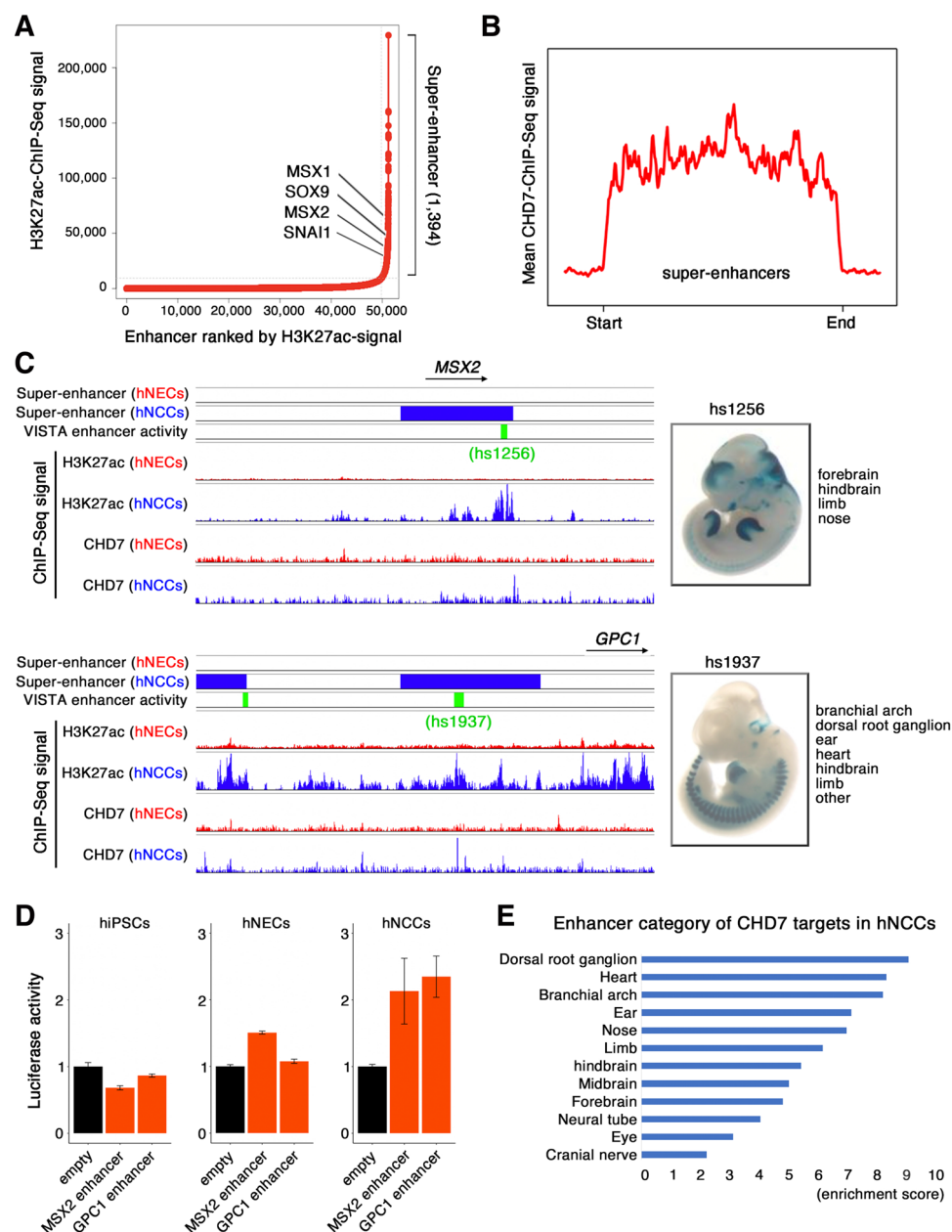
**Co-occupancy of CHD7 with TFAP2A and nuclear receptor subfamily 2 group F1/F2 (NR2F1/F2).** The master NC regulators TFAP2A and NR2F1/F2 have been reported to be involved in the hNCC enhancer regulatory mechanism<sup>38</sup>. Given that motif analysis of the CHD7-binding region in hNCCs revealed the possible co-existence of TFAP2A with CHD7 (Fig. 2C), we analyzed the relationship among CHD7, TFAP2A, and NR2F1/F2 in hNCCs. Reanalysis of published datasets for TFAP2A, NR2F1, and NR2F2 ChIP-seq experiments in hNCCs<sup>38</sup> showed that TFAP2A bound preferentially to the proximal region of its targets. Addition-



**Figure 2.** Cell type-specific functions of CHD7. **(A)** Heatmap analyses of ChIP-seq signal of EP300 and selected histone modifications in hiPSCs, hNECs, and hNCCs (H3K4me1, H3K4me3, and H3K27ac). Read counts were extracted for all ChIP-seq experiments within a region spanning  $\pm 5$  kb around the center of each annotated CHD7-peaks detected in each cell type (1st row: 31,518 peaks in hiPSCs, 2nd row: 17,941 peaks in hNECs and 3rd row: 5,357 peaks in hNCCs). **(B)** Overlaps of CHD7-binding regions across different cell types. **(C)** Cell type-specific CHD7-binding sequences were analyzed using the top 500 binding regions of CHD7-ChIP-seq data. The detected sequences were compared using the JASPAR motif database. **(D)** GO analysis for biological processes associated with CHD7-bound regions in hiPSCs, hNECs, and hNCCs. Bar plots displaying the top 10 GO categories.

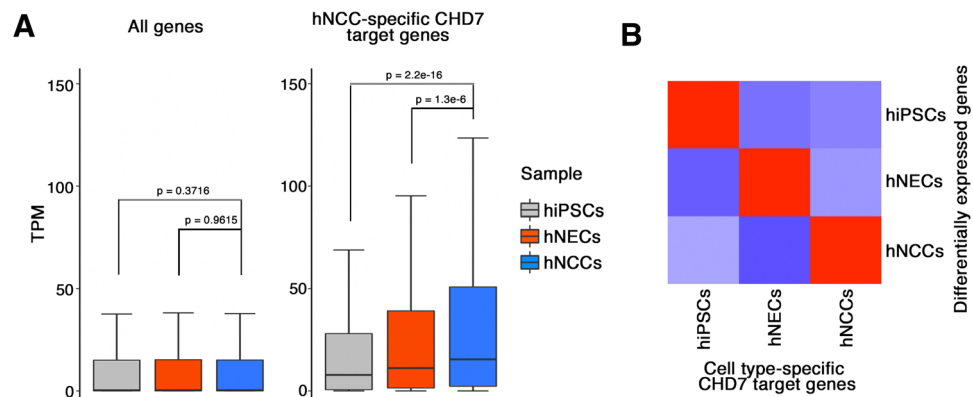
ally, its bound regions were associated with H3K4me3 and EP300, indicative of the role of TFAP2A as a direct transcriptional activator (Supplementary Fig. S2). In contrast, the NF2F1/F2-binding datasets showed that a substantial fraction of NR2F1/2 bound to the distal regions of their targets with increased enhancer activity,





**Figure 3.** ChD7 bound to SE regions with TFAP2A and NR2F1/2. **(A)** Distribution of H3K27ac ChIP-seq density at enhancers in increasing order of input-normalized H3K27ac ChIP-seq signal. Enhancers above the curve's inflection point are considered SEs. The respective ranks of SEs and their associated genes are shown. **(B)** Density plots of mean ChD7 ChIP-seq signals across SEs in hNCCs. **(C)** Snapshots showing combined tracks of SEs (red; hNEC, blue; hNCC), VISTA enhancer activity (green), and ChIP-seq data (α-H3K27ac and α-ChD7) near *MSX2* and *GPC1*. Images on right are screenshots of tissue-specific activities of the enhancer proximal to *MSX2* and *GPC1* tested in a lacZ reporter transgenic mouse assay extracted from the VISTA Enhancer Browser database (<https://enhancer.lbl.gov/>). **(D)** A luciferase reporter assay was performed to determine the activity of ChD7-bound enhancers near *MSX2* and *GPC1*. Reporter plasmid (empty control, *MSX2* enhancer, and *GPC1* enhancer) was transfected to hiPSCs, hNECs, and hNCCs to compare enhancer activity. Renilla reporter was co-transfected for normalization (F/R ratio). Y axis shows fold enrichment over an empty control. Data are presented as the mean ± SEM. **(E)** Functional analysis of genomic regions of ChD7-bound regions intersected with enhancer regions validated in the VISTA Enhancer Browser database in hNCCs. Bar plot showing enrichment of enhancer occupancy across 16 different tissues.

demarcated by co-occupancy of H3K4me1 and H3K27ac (Supplementary Fig. S2). We investigated the interplay of ChD7 with TFAP2A and NR2F1/2 in hNCCs by examining the enrichment of TFAP2A and NR2F1/2 in



**Figure 4.** Correlation of CHD7-bound enhancer regions with transcription of target genes in hNCCs. **(A)** Boxplot showing expression of all genes (left) and hNCC-specific CHD7-target genes (right) in hiPSCs, hNECs, and hNCCs. **(B)** Correlation analysis of cell type-specific CHD7 target genes and DEGs in hiPSCs, hNECs, and hNCCs. Heatmap showing z-score of correlations.

the CHD7-binding region in hNCCs. As a reference, we also compared their enrichment in the CHD7-binding region in hNECs. Interestingly, these transcription factors were preferentially enriched in hNCC-specific CHD7-binding regions, and their abundance was substantially lower in hNEC-specific CHD7-binding regions (Fig. 5A–C). These results indicate that the co-existence of tissue-specific transcription factors ensures cell type-specific CHD7 binding. A previous study<sup>38</sup> showed that hNCC-specific gene expression was linked to TFAP2A- or NR2F1/2-binding regions. In particular, it has been shown that TFAP2A and NR2F1/2 are present in the enhancer regions of hNCC genes such as *SOX9*, *SNAIL1*, and *TWIST1*, which are essential genes for neural crest induction and differentiation<sup>39</sup>. Notably, these genes are also targeted by CHD7 in hNCCs (Fig. 5D). To directly examine whether CHD7 actually interacts with TFAP2A and NR2F1/2 in hNCC-specific CHD7 targets, we performed sequential ChIP experiments called ChIP-reChIP experiments<sup>40</sup> (Fig. 5D). ChIP-ReChIP-qPCR analysis showed enrichment of CHD7-TFAP2A, CHD7-NR2F1, and CHD7-NR2F2 signals over the CHD7-IgG signal at enhancers, but not at control regions (Fig. 5D). These observations indicated co-occupancy of CHD7, TFAP2A, and NR2F1/2 on CHD7-targeted regions in hNCC. Furthermore, to investigate the impact of the co-occupancy of these factors on their targets in hNCCs, we examined the gene expression classified by the degree of binding of each factor (Fig. 5E). Interestingly, genes co-occupied by CHD7 together with TFAP2A and NR2F1/2 further elaborated hNCC-specific gene expression, indicating the synergistic actions of these factors as a basis for cell type-specific transcription in hNCCs (Fig. 5E and Supplementary Fig. S4A).

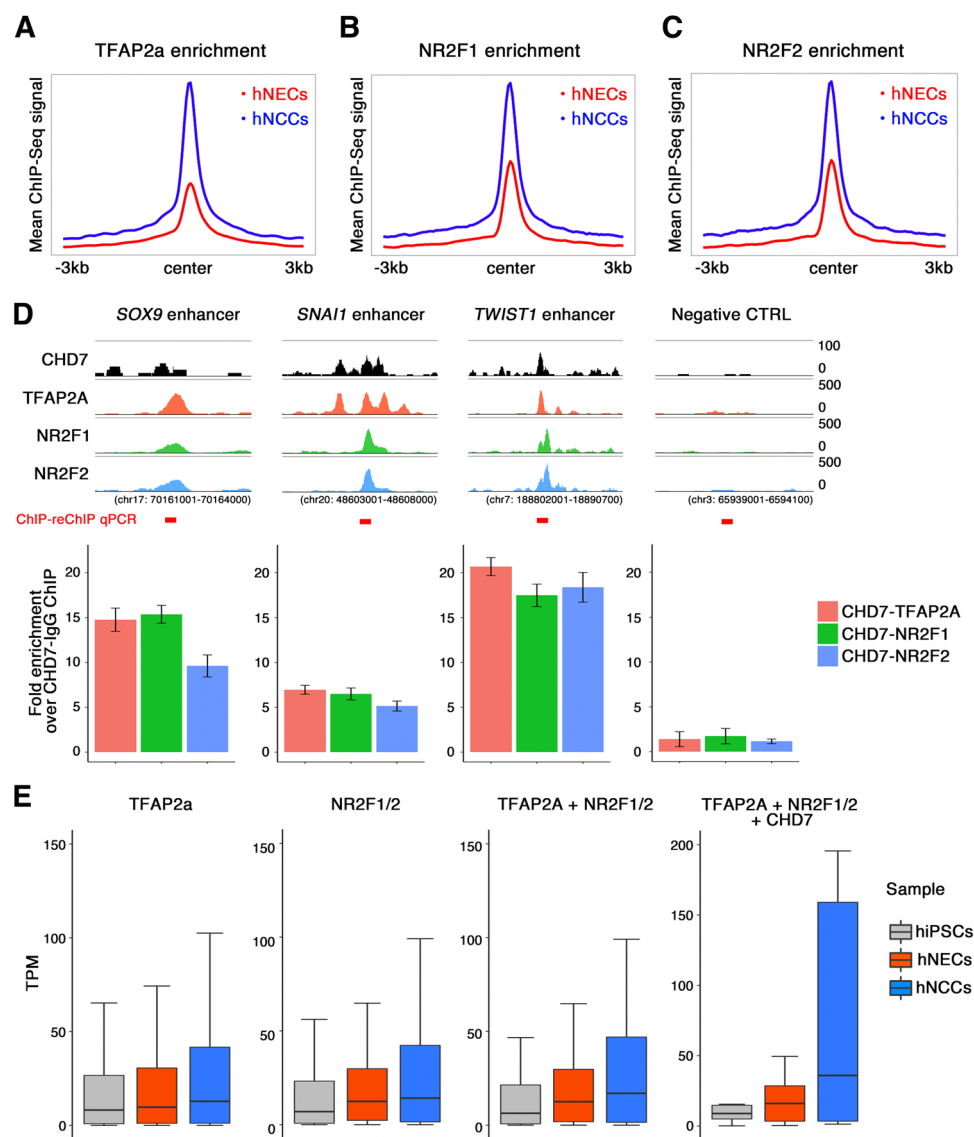
## Discussion

This study provides the first genome-wide mapping of CHD7 in hNCCs. Combined with chromatin-state profiling, it revealed that CHD7 acts as an activator of enhancers to specify hNCC-specific gene expression. Importantly, cell type-specific CHD7-binding sites in hNCCs were enriched in SEs and predicted critical transcription factor-binding regions, such as TFAP2A, that drive hNCC-specific cell state-determining gene expression. Furthermore, the presence of CHD7 in hNCC transcription factor-binding regions increased the accuracy of detecting genes with hNCC-specific expression patterns, suggesting that profiling of CHD7-binding properties in hNCCs may provide an excellent resource for hNCC biology.

NCC populations arise transiently from the neural tube during development<sup>41–43</sup>. Importantly, they acquire broad differentiation and migration potentials and contribute to craniofacial bone, cartilage, peripheral nervous system, pigment cells, and specific heart structures<sup>41–43</sup>. Because NCCs contribute to the construction of the cranial mesenchyme, abnormal NC development is associated with various congenital malformations known as neurocristopathies<sup>44</sup>, which manifest as hearing loss and complex craniofacial defects and comprise one of the most common types of congenital disabilities<sup>45</sup>. Because the NC cannot be studied in the embryonic context, NCC induction systems using PSCs are essential for investigating the molecular mechanisms regulating the cellular behavior of NCCs, especially in humans<sup>11,12</sup>.

Based on this background, unbiased high-throughput transcriptomic and epigenomic profiling of hNCCs can be performed to identify cis- and trans-regulatory factors that control hNCC development<sup>38,39,46,47</sup>. The master lineage designator TFAP2A and the nuclear receptor NR2F1/2 simultaneously recognize the enhancer-coded sequence in hNCCs, leading to a transcriptionally permissive enhancer chromatin state in hNCCs<sup>39</sup>. The current analysis showed that these factors cooperated with CHD7 in hNCCs and that the regions co-occupied with TFAP2A, NR2F1/2, and CHD7 were highly linked to hNCC-specific gene expression (Fig. 5 and Supplementary Fig. S3). The presence or absence of CHD7 binding may thus predict the cell-specific genes essential for development and cell function in hNCCs.

We previously demonstrated the defective migration of iPSC-NCCs derived from CHARGE patients (CHARGE-iPSC-NCCs)<sup>12</sup>, and also reported 338 differentially expressed genes (DEGs) (238 upregulated and 100 downregulated) in CHARGE-iPSC-NCCs<sup>12</sup>. We therefore considered that it would be interesting to determine if the datasets obtained in the current study could help to elucidate the molecular mechanism underlying



**Figure 5.** Co-localization of CHD7, TFAP2A, and NR2F1/2 strongly induced hNCC-specific gene expression. (A) Enrichment analysis of TFAP2A in hNECs and hNCCs. The mean TFAP2A ChIP-seq signal from  $\pm 3$  kb of the CHD7-bound center was plotted. (B) Enrichment analysis of NR2F1 around CHD7-bound regions in hNECs and hNCCs. The mean NR2F1-ChIP-seq signal was plotted. (C) Enrichment analysis of NR2F2 around CHD7-bound regions in hNECs and hNCCs. The mean NR2F2-ChIP-seq signal was plotted. (D) ChIP-reChIP analysis of CHD7-TFAP2A-NR2F1/2 co-occupancy. Snapshot of ChIP-Seq signal (CHD7: black, TFAP2A: red, NR2F1: green, NR2F2: blue) in hNCC enhancer region of NCC specific genes. The genomic region indicated the whole loci shown in the snapshot. The red bar indicates the amplified region at ChIP-reChIP qPCR analysis. ChIP-reChIP analysis was performed using an anti-CHD7 antibody as the first antibody. Antibodies against control IgG, TFAP2A (red), NR2F1 (blue), or NR2F2 (green) antibodies were used for reChIP, respectively. Immunoprecipitated DNA was analyzed by qPCR of selected hNCC enhancers shown in the upper panels. Y axis shows fold enrichment over a CHD7-IgG ChIP-reChIP. (E) Target genes of CHD7, TFAP2A, and NR2F1/2 were extracted and gene expressions in each cell type were plotted.

the cellular phenotypes of CHARGE-iPSC-NCCs. We performed analyses to integrate these data and found that, although CHD7-bound genes were expected to be downregulated in CHARGE syndrome, the downregulated DEGs were not included in the CHD7-bound gene lists, while about ten upregulated DEGs in CHARGE-iPSC-NCCs were included as CHD7-bound genes (data not shown).

Furthermore, we also examined the expression of genes bound by CHD7, TFAP2A, and NR2F1/2. Although a subset of genes was slightly downregulated in CHARGE-iPSC-NCCs, there was no major difference in the expression of these target genes (Supplementary Fig. S4). Further analyses should be performed to investigate this divergence in results in CHARGE-iPSC-NCCs.

One potential limitation of the current study was the axial-specific heterogeneity of the hNCCs. NC development progresses from anterior to posterior within the embryo during neurulation<sup>41,42,48</sup>, and despite their multipotency, NCCs produce different derivatives depending on their axial position<sup>48</sup>. Given that the hNCCs used in the current study were cranial NCCs<sup>11,12</sup>, and CHD7 may have different targets in other types of NCCs. Furthermore, this analysis compares multiple ChIP-seq data published by different groups using different antibodies. Therefore, it may have affected peak detection compared to the analysis that directly compares the signal intensity between each cell type.

Non-coding genomic mutations are implicated in human diseases, with associations between mutations in the enhancer region and the pathogenesis of disease progression<sup>47,49,50</sup>. Furthermore, mutations in tissue-specific enhancer regions can disrupt target-gene regulation in specific tissues. Indeed, mutations in the enhancer region of SOX9, an essential gene for NC development, have also been reported in the human craniofacial disorder Pierre Robin sequence<sup>43</sup>. Given that the impact of mutations in non-coding areas on long-term functional outcomes is difficult to address, the current results may help to determine the significance of genetic mutations in the enhancer regions in diseases considered to be caused by defects in NC development.

## Methods

**Cell culture.** The hiPSC line 1210B2 was maintained as described previously<sup>51,52</sup>. The hiPSCs were dissociated using 0.5× TrypLE Select (Thermo Fisher) and plated on iMatrix511 (Nippi)-coated dishes in StemFit (AK03N) medium at feeder-free condition. The hiPSC-derived NECs were established and maintained as described previously<sup>53,54</sup>. The hNECs were passaged every 4 days and plated at a ratio of 1:5. The cells were dissociated using TrypLE Select (Thermo Fisher) and plated on Matrigel (Corning)-coated dishes in RHB-A medium (Takara) supplemented with 10 ng/mL EGF (Peprotech) and 10 ng/mL FGF2 (Peprotech). The hiPSC-derived NCCs were established previously and passaged less than three times for analysis. The hNCC medium consisted of 1:1 neurobasal medium (Thermo Fisher Scientific) and D-MEM/Ham's F-12 (Wako) medium containing 1× GlutaMax (Thermo Fisher Scientific), 0.5× GEM 21 NeuroPlex serum-free supplement (Gemini Bio Products), 0.5× N2 supplement (Thermo Fisher Scientific), 5 mg/mL insulin (Sigma-Aldrich), 0.5% penicillin and streptomycin (Nacalai tesque), 20 ng/mL FGF2 and 20 ng/mL EGF. The hNCCs were dissociated using TrypLE Select and plated on fibronectin (Sigma-Aldrich)-coated dishes.

**Plasmid construction.** Selected genomic regions near the human MSX2 and GPC1 genes were amplified by PCR from the human genomic DNA and cloned into a pGL3 promoter vector (Promega). Primers used for cloning were listed in Supplemental Table S4.

**Luciferase assay.** hiPSCs, hNECs, and hNCCs were replated onto 12-well plates (Greiner Bio-One) and transfected with a reporter vector the next day. The sea pansy luciferase gene conjugated with human elongation factor-1α promoter (R-Luc) was also cotransfected as an internal control. Transfection was performed using a genejuice (Merck) according to the manufacturer's procedures. Cells were solubilized 24 h after the transfection, and luciferase activity was measured according to the procedures recommended for the Dual-Luciferase Reporter Assay System (Promega) using GLOMAX multi-detection system (Promega). Firefly luciferase activity was divided by R-Luc activity from the cotransfected control plasmid.

**ChIP-reChIP qPCR.** Chromatin immunoprecipitation was performed as previously described with minor modifications<sup>10,39</sup>. The dual cross-linking method was performed using 200 mM EGS for 30 min followed by 1% formaldehyde for 10 min and quenched by adding glycine to a final concentration of 125 mM. Cells were collected with NP-40 buffer (10 mM Tris-HCl, 10 mM NaCl, 0.5% NP-40) and centrifuged to a pellet, lysed in SDS lysis buffer (50 mM Tris-HCl, 1% SDS, 10 mM EDTA), and then incubated for 10 min on ice. Cell lysates were sonicated using Covaris S2 (Covaris) until the DNA fragments were 200–300 base pairs in length. These chromatin samples were diluted 1:10 with dilution buffer [50 mM Tris-HCl, 167 mM NaCl, 1.1% Triton X-100, 0.11% sodium deoxycholate (NaDOC)], and 5% of the total volume was stored as input. The first immunoprecipitation was performed at 4 °C overnight with 3 μg of an antibody conjugated with Dynabeads (VERITAS). Immune complexes were washed sequentially with RIPA low-salt (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, 0.1% NaDOC), RIPA high-salt (50 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% Triton X-100, 0.1% NaDOC), TE and eluted by TE/10 mM DTT at 37 °C for 30 min. Eluted chromatin was diluted 20 times with RIPA low-salt buffer for the second immunoprecipitation reaction. Next, immune complexes were washed, and immunoprecipitated DNA was eluted with elution buffer (10 mM Tris-HCl, 300 mM NaCl, 5 mM EDTA, 0.5% SDS) at 65 °C for 4 h. DNA was further incubated with RNase and proteinase K (Nacalai Tesque) and purified using a ChIP DNA clean and concentrator kit (Zymo Research). Antibodies used in this study have been reported as suitable for ChIP assay: anti-CHD7 (D3F5, Cell Signaling Technology), anti-TFAP2A (sc-12726, Santa Cruz Biotechnology), anti-NR2F1 (PP-H8132-00, Perseus Proteomics) and anti-NR2F2 (PP-H8132-00, Perseus Proteomics). Quantitative PCR (qPCR) analysis was performed using a ViiA7 system (Applied Biosystems) with KAPA SYBR FAST Master Mix (KAPA Biosystems). The primers used are listed in Supplemental Table S4.

**Reanalysis of ChIP-seq data.** ChIP-seq data for histone modification in hNCCs were obtained from the Gene Expression Omnibus (GEO) database (GSE28876). Other ChIP-seq data for CHD7-binding in hNCCs and other ChIP-seq data for hiPSCs and hNECs were deposited as a part of our previous study through GEO accession number GSE108506. Raw reads were trimmed based on read length and quality using Trimmomatic (v. 0.33). The trimmed reads were then aligned to the reference genome (UCSC hg19) using Bowtie2 (v. 2.1.0) with



the default parameters, and only uniquely mapped reads were used for subsequent analyses. The resulting SMA files were converted to the BAM format using SAMtools (v. 0.1.19). Peak calling was performed using MACS2 (v. 2.2.7.1) with default settings<sup>55</sup>. Cell type-specific CHD7-peaks were identified using the DiffBind (v. 3.6.1) R package. Overlaps for each peak were calculated using bedtools (v. 2.17.0) with default parameters. The genome-wide peak distribution from TSS was calculated using the ChIPseeker (v. 1.32.0) R package, and functional GO analysis of CHD7 peaks was performed using the GREAT website (<http://bejerano.stanford.edu/great/public/html/index.php>). The enhancer category was calculated based on the overlap of the CHD7-binding site and the enhancer region registered in the VISTA enhancer browser.

**Heatmap analysis.** Heatmaps for each sample against the CHD7-binding region were made using deepTools (v. 1.5.11). The average H3K27ac read density at each region and the corresponding flanking region were calculated (bin size = 50). The lengths of the SE regions (between TSS and TES) were scaled relative to the median length.

**Motif search.** Enriched sequences in the top 500 CHD7-binding regions were searched using MEME (v. 1.4) software and the identified sequences were compared using motif databases (JASPAR and HOCOMOCO) with TOMTOM (v. 2.2) software.

**Identification of SEs 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 for H3K27ac<sup>31</sup>. We used a promoter-exclusion zone of 5000 bp to prevent any enhancer contained within a window of  $\pm 2500$  bp around an annotated TSS from being stitched. Enhancer-associated genes were defined based on the calculated distance from the center of the SE to the nearest TSS of each gene. The assignment of enhancers to the closest genes in hNCCs is shown in Supplementary Table S2.

**Library preparation and RNA-seq.** Samples for RNA-seq were prepared using the TruSeq RNA Library Prep Kit (Illumina) according to the manufacturer's protocol. The sequencing library was sequenced on a Next-Seq 1000 (Illumina). The adapter sequence and low-quality reads were removed using trimmomatic (v. 0.33), and the trimmed reads were mapped to reference genome hg38 using Kallisto (v. 0.46.2) with default settings. The count matrix data from Kallisto was loaded into R software (v. 4.2.0), and downstream analysis was performed. Differentially expressed genes were evaluated using the limma (v. 3.52.2) library.

**Computational analysis of CHARGE patient microarray data.** Our previously published gene expression profiles of WT and CHARGE patient-derived NCCs data (GSE86212) were loaded into R (v. 4.2.0) for statistical analysis. Boxplots were made using the R package ggplot2 (v. 3.3.6).

## Data availability

RNA-seq data have been deposited in the NCBI GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) and are accessible through the following GEO accession number (GSE220680).

Received: 13 September 2022; Accepted: 29 December 2022

Published online: 31 December 2022

## References

- Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
- Li, W., Notani, D. & Rosenfeld, M. G. Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. *Nat. Rev. Genet.* **17**, 207–223 (2016).
- West, J. A. *et al.* Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming. *Nat. Commun.* **5**, 4719 (2014).
- Hu, Z. & Tee, W. W. Enhancers and chromatin structures: Regulatory hubs in gene expression and diseases. *Biosci. Rep.* **37**, 20160183 (2017).
- Zollino, M. *et al.* Syndromic craniosynostosis can define new candidate genes for suture development or result from the non-specific effects of pleiotropic genes: Rasopathies and chromatinopathies as examples. *Front. Neurosci.* **11**, 587 (2017).
- Ciptasari, U. & van Bokhoven, H. The phenomenal epigenome in neurodevelopmental disorders. *Hum. Mol. Genet.* **29**, R42–R50 (2020).
- Vissers, L. E. *et al.* Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.* **36**, 955–957 (2004).
- Blake, K. D. & Prasad, C. CHARGE syndrome. *Orphanet J. Rare Dis.* **1**, 34 (2006).
- Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: Advances through genomics. *Nat. Rev. Genet.* **10**, 161–172 (2009).
- Chai, M. *et al.* Chromatin remodeler CHD7 regulates the stem cell identity of human neural progenitors. *Genes Dev.* **32**, 165–180 (2018).
- Bajpai, R. *et al.* CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature* **463**, 958–962 (2010).
- Okuno, H. *et al.* CHARGE syndrome modeling using patient-iPSCs reveals defective migration of neural crest cells harboring CHD7 mutations. *Elife* **6**, 21114 (2017).
- Whittaker, D. E. *et al.* The chromatin remodeling factor CHD7 controls cerebellar development by regulating reelin expression. *J. Clin. Invest.* **127**, 874–887 (2017).
- Schnetz, M. P. *et al.* Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. *Genome Res.* **19**, 590–601 (2009).

15. Micucci, J. A. *et al.* CHD7 and retinoic acid signaling cooperate to regulate neural stem cell and inner ear development in mouse models of CHARGE syndrome. *Hum. Mol. Genet.* **23**, 434–448 (2014).
16. He, D. *et al.* Chd7 cooperates with Sox10 and regulates the onset of CNS myelination and remyelination. *Nat. Neurosci.* **19**, 678–689 (2016).
17. Feng, W. *et al.* Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. *Nat. Commun.* **8**, 14758 (2017).
18. Liu, C. *et al.* CHD7 regulates bone-fat balance by suppressing PPAR- $\gamma$  signaling. *Nat. Commun.* **13**, 1989 (2022).
19. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
20. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
21. Nitsch, S., Zorro Shahidian, L. & Schneider, R. Histone acylations and chromatin dynamics: Concepts, challenges, and links to metabolism. *EMBO Rep.* **22**, e52774 (2021).
22. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).
23. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Genome Biol.* **8**, R24 (2007).
24. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
25. Siegrist, C. A. *et al.* RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer. *Mol. Cell Biol.* **13**, 6375–6384 (1993).
26. de Croze, N., Maczkowiak, F. & Monsoro-Burq, A. H. Reiterative AP2a activity controls sequential steps in the neural crest gene regulatory network. *Proc. Natl. Acad. Sci. USA* **108**, 155–160 (2011).
27. Rothstein, M. & Simoes-Costa, M. Heterodimerization of TFAP2 pioneer factors drives epigenomic remodeling during neural crest specification. *Genome Res.* **30**, 35–48 (2020).
28. Schulz, Y. *et al.* CHD7, the gene mutated in CHARGE syndrome, regulates genes involved in neural crest cell guidance. *Hum. Genet.* **133**, 997–1009 (2014).
29. Yan, S. *et al.* CHD7 regulates cardiovascular development through ATP-dependent and -independent activities. *Proc. Natl. Acad. Sci. USA* **117**, 28847–28858 (2020).
30. Loven, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
31. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
32. Cheung, M. & Briscoe, J. Neural crest development is regulated by the transcription factor Sox9. *Development* **130**, 5681–5693 (2003).
33. Tribulo, C., Aybar, M. J., Nguyen, V. H., Mullins, M. C. & Mayor, R. Regulation of Msx genes by a Bmp gradient is essential for neural crest specification. *Development* **130**, 6441–6452 (2003).
34. Monsoro-Burq, A. H., Wang, E. & Harland, R. Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during *Xenopus* neural crest induction. *Dev. Cell* **8**, 167–178 (2005).
35. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA enhancer browser—A database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88–92 (2007).
36. Catron, K. M., Wang, H., Hu, G., Shen, M. M. & Abate-Shen, C. Comparison of MSX-1 and MSX-2 suggests a molecular basis for functional redundancy. *Mech. Dev.* **55**, 185–199 (1996).
37. David, G. *et al.* Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. *J. Cell Biol.* **111**, 3165–3176 (1990).
38. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
39. Rada-Iglesias, A. *et al.* Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* **11**, 633–648 (2012).
40. Furlan-Magaril, M., Rincon-Arango, H. & Recillas-Targa, F. Sequential chromatin immunoprecipitation protocol: ChIP-reChIP. *Methods Mol. Biol.* **543**, 253–266 (2009).
41. Couly, G. F., Coltey, P. M. & Le Douarin, N. M. The triple origin of skull in higher vertebrates: A study in quail-chick chimeras. *Development* **117**, 409–429 (1993).
42. Bronner, M. E. & LeDouarin, N. M. Development and evolution of the neural crest: An overview. *Dev. Biol.* **366**, 2–9 (2012).
43. Le Douarin, N. M. & Dupin, E. The “beginnings” of the neural crest. *Dev. Biol.* **444**(Suppl 1), S3–S13 (2018).
44. Benish, B. M. Letter: “The neurocristopathies: A unifying concept of disease arising in neural crest development”. *Hum. Pathol.* **6**, 128 (1975).
45. Passos-Bueno, M. R., Ornelas, C. C. & Fanganiello, R. D. Syndromes of the first and second pharyngeal arches: A review. *Am. J. Med. Genet. A* **149A**, 1853–1859 (2009).
46. Prescott, S. L. *et al.* Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83 (2015).
47. Long, H. K. *et al.* Loss of extreme long-range enhancers in human neural crest drives a craniofacial disorder. *Cell Stem Cell* **27**, 765–783 (2020).
48. Dupin, E., Creuzet, S. & Le Douarin, N. M. The contribution of the neural crest to the vertebrate body. *Adv. Exp. Med. Biol.* **589**, 96–119 (2006).
49. Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).
50. Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
51. Nakagawa, M. *et al.* A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. *Sci. Rep.* **4**, 3594 (2014).
52. Okita, K. *et al.* 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 (2013).
53. Falk, A. *et al.* 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 (2012).
54. Isoda, M. *et al.* 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 (2016).
55. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

## Acknowledgements

We are very grateful to the members of the Department of Physiology, Keio University School of Medicine. We also thank E. Nishi, R. Shimamura, I. Koya, and M. Chai for their assistance with the experiments, and Dr. Shinya Yamanaka (Kyoto University) for providing the hiPSC clone (1210B2). This work was supported by JSPS

KAKENHI Grant Number JP21K06396 to T.S. and JP19H03623 to J.K. We also thank Susan Furness, Ph.D., from Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript.

### Author contributions

Conceptualization, T.S., J.K.; methodology, T.S., H.Oku., N.M., T.A.-N., M.S., R.T., S.B.; investigation, T.S., N.M., J.K.; writing the original draft, T.S., J.K.; writing, reviewing, and editing, H.Oka.; funding acquisition, T.S., J.K.

### Competing interests

H.O. is a compensated scientific consultant of San Bio, Co., Ltd., and K Pharma Inc. The other authors declare no potential conflicts of interest.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-27293-6>.

**Correspondence** and requests for materials should be addressed to J.K. or H.O.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022