NEUROD1 Intrinsically Initiates Differentiation of Induced Pluripotent Stem Cells into Neural Progenitor Cells

Won-Young Choi[1], Ji-Hyun Hwang[1], Ann-Na Cho[2], Andrew J. Lee[3], Inkyung Jung[3], Seung-Woo Cho[2], Lark Kyun Kim[4],*, and Young-Joon Kim[5],*

[1]Interdisciplinary Program of Integrated OMICS for Biomedical Science, The Graduate School, Yonsei University, Seoul 03722, Korea, [2]Department of Biotechnology, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Korea, [3]Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea, [4]Severance Biomedical Science Institute and BK21 PLUS Project for Medical Sciences, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul 06273, Korea, [5]Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 03722, Korea
*Correspondence: LKKIM@yuhs.ac (LKK); yjkim@yonsei.ac.kr (YJK)
https://doi.org/10.14348/molcells.2020.0207
www.molcells.org

Cell type specification is a delicate biological event in which every step is under tight regulation. From a molecular point of view, cell fate commitment begins with chromatin alteration, which kickstarts lineage-determining factors to initiate a series of genes required for cell specification. Several important neuronal differentiation factors have been identified from ectopic over-expression studies. However, there is scarce information on which DNA regions are modified during induced pluripotent stem cell (iPSC) to neuronal progenitor cell (NPC) differentiation, the *cis* regulatory factors that attach to these accessible regions, or the genes that are initially expressed. In this study, we identified the DNA accessible regions of iPSCs and NPCs via the Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq). We identified which chromatin regions were modified after neuronal differentiation and found that the enhancer regions had more active histone modification changes than the promoters. Through motif enrichment analysis, we found that NEUROD1 controls iPSC differentiation to NPC by binding to the accessible regions of enhancers in cooperation with other factors such as the Hox proteins. Finally, by using Hi-C data, we categorized the genes that directly interacted with the enhancers under the control of NEUROD1 during iPSC to NPC differentiation.

Keywords: chromatin accessibility, Hi-C, induced pluripotent stem cell, NEUROD1, neuronal progenitor cell

INTRODUCTION

Human induced pluripotent stem cells (hiPSCs) have been heavily utilized as experimental platforms by various research areas studying lineage differentiation and modelling of human developmental process and diseases (Cotovio and Fernandes, 2020; Rowe and Daley, 2019; Shum et al., 2015). The need for iPSC is especially high in disease modelling given the difficulties in isolating cells from patients and obtaining the optimal number of cells for experiments. Among different cell lineages from iPSC, neuronal lineage which can be matured to functional neuron is of special clinical interest, and iPSC-derived neuronal differentiation models are already

in use in analysis of neurodegenerative disease such as Alzheimer's disease, Parkinson's disease, Huntington's disease and motor neuron diseases (Chang et al., 2020; Wu et al., 2019). Although iPSC-derived NPC model has been well studied in terms of defining cellular phenotype and neuronal function according to gene expression, the mechanism of gene regulation is still elusive (Kang et al., 2017; Zhang et al., 2013).

Gene expression is tightly regulated during cell fate commitment (Rue and Martinez Arias, 2015), and one extensively studied layer of regulation is epigenetic modification such as histone and DNA methylation, which regulates the nuclear environments to control gene transcription (Cedar and Bergman, 2009). Other than direct modification, chromatin accessibility to nuclear macromolecules such as transcription factors (TFs) itself is another facet in gene regulation (Klemm et al., 2019). Highly accessible regions of chromatin are considered to be open and promote gene expression, whereas inaccessible regions are closed and inhibit expression (Cedar and Bergman, 2009). It is therefore important to localize the regions with cell-type specific accessibilities as they are inferred to contain genes that determine cell fate (Scott-Browne et al., 2016). Enhancers, for example, have been identified as regions that hold most chromatin accessibiity changes during neuronal development (Frank et al., 2015).

The enhancer is one of the *cis* regulatory elements known to regulate gene expression via physical interaction with target genes from distal genomic regions (Schaffner, 2015). At the point of cell fate transition, lineage-specific enhancers initiate the first wave of gene expression in collaboration with lineage-determining TFs. First discovered and studied in Drosophila, these basic helix-loop-helix (bHLH) TFs are critical in neuronal differentiation, and have been coined proneural factors to emphasize the importance of their role (Ross et al., 2003). Similar studies in mammals initially showed that that a trio of proneural or BAM factors (BRN2 [POU3F2], ASCL1, and MYT1L) initiate changes in chromatin accessibility during the direct reprogramming of dermal fibroblasts into neuronal cells (Wapinski et al., 2013). Further research indicated that NEUROD1 is another proneural factor that promotes neuronal development in a synergistic manner with the BAM factors and is solely sufficient for direct neuronal reprogramming (Pang et al., 2011; Pataskar et al., 2016). However, following the characterization of initial TFs involved in neuronal differentiation and discovery of some target enhancer regions for BAM factors and NEUROD1 by overexpressing the proneural factors in dermal fibroblasts (Wapinski et al., 2013), much still remains unknown about which DNA regions undergo chromatin accessibility changes, where the enhancers for neuronal development are, or which factors can bind to the enhancers during stem cell differentiation.

In this study, we identified cell type-specific accessible chromatin regions of iPSCs and NPCs by Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) and also noted the DNA regions that changed during differentiation and found that the accessibility of the enhancer, but not the promoter, was related to active histone modification changes. Through motif enrichment analysis, we found that NEUROD1, Hox proteins, and ONECUT1 may control neuro-nal differentiation by binding to the accessible regions of the enhancer. Furthermore, we identified the initial genes that are controlled by NEUROD1 during iPSC to NPC differentiation using Hi-C sequencing. Finally, our results can be used as additional resource to aid studies of regulatory mechanism in neuronal development-related genes during iPSC to NPC differentiation by chromatin accessibility changes and TF bindings.

MATERIALS AND METHODS

Culture for iPSC cells and iPSC-derived neural progenitor cells
The use of human iPSCs was approved by the Institutional Review Board (IRB) of Yonsei University (permit No. 7001988-201802-BR-119-01E). In this study, we used pre-constructed iPSC cell lines (Cat. No. C-004-5C; Gibco, USA). We cultured iPSC cell lines on Matrigel (BD Biosciences, USA) with Essential 8 medium (Invitrogen, USA) coating (Chen et al., 2011; 2017; Higuchi et al., 2015). To differentiate iPSC to NPC, an embryoid body (EB) was generated through culturing of human iPSCs for 5-6 days on non-adherent petri dishes in Essential 8 (Invitrogen) with additional supplement of 5 μM dorsomorphin (DM; Sigma-Aldrich, USA) and 5 μM SB431542 (Sigma-Aldrich). EBs were then attached to the new culture dish coated with Matrigel (BD Biosciences) with neural induction medium comprised of DMEM/F12 medium (Invitrogen), 1× N2 supplement (Invitrogen), 1× nonessential amino acids (Invitrogen) for 6 days. When neural rosette formation appears in the center of the EBs, NPCs were collected by using Pasteur pipet (Jin et al., 2019; Seo et al., 2015; Shin et al., 2017).

ATAC-seq
ATAC-seq samples were prepared by described method (Buenrostro et al., 2015). We washed 50,000 cells with cold phosphate-buffered saline (PBS) and resuspended cells by cell lysis buffer (10 mM Tris-HCl [pH 7.4], 10 mM NaCl, 3 mM MgCl$_2$, 0.1% IGEPAL CA-630) and centrifuged cells (10 min, 2,000g at 4°C). After centrifugation, the nuclei pellet were resuspended with Transposase mixture (12.5 μl 2× TD buffer, 12 μl Nuclease-free water and 0.5 μl of Tn5 per sample) in Illumina Nextera Kit (FC-121-1031; Illumina, USA) and incubated for 30 min at 37°C. Tagmented DNA fragments were puri ed using MinElute PCR purification kit (Qiagen, Germany) following manufacturer's protocol. For eluted DNA samples, we executed small scale quantitative PCR (qPCR) to determine adequate PCR amplification cycles for library preparation. We checked Ct value for qPCR sample mix (5 μl 2× KAPA HiFi HotStart Ready Mix, 0.1 μl 100× SYBR green, 0.3 μl Nextera Ad1.2 [25 μM], 0.3 μl Nextera Ad2.X [25 μM], 2.3 μl Nuclease-free water and 2 μl eluted DNA) and determined PCR amplification cycles with value of Ct+1. We then executed large scale PCR (25 μl 2× KAPA HiFi HotStart Ready Mix, 1.5 μl Nextera Ad1.2 [25 μM], 1.5 μl Nextera Ad2.X [25 μM], 2 μl Nuclease-free water and 20 μl eluted DNA). After PCR step, we cleaned library samples and performed size selection (200-400 bp) by AMPure XP beads (A63881; Beckman Coulter, USA) following manufacturer's instruction. The

quality of libraries was checked by Agilent Technologies 2100 Bioanalyzer (Agilent, USA) regarding concentration and size. Finally, we sequenced libraries with Illumina HiSeq 2500 with 101 bp paired-end reads.

ATAC-seq data processing and Stage-specific peak calling
Prior to mapping, quality control steps were executed. We assessed reads by 'FastQC' for read quality control and adapter removing process was done by 'Cutadapt' (cutadapt -a CTGTCTCTTATACACATCT -q 10 --minimum-length 36 -o sample_1_trimmed.fastq -p sample_2_trimmed.fastq sample_1.fastq sample_2.fastq). Processed reads were aligned by 'Bowtie2' (bowtie2 --very-sensitive -x <genomeIndexName> -X 2000 -1 sample_1_trimmed.fastq -2 sample_2_trimeed. fastq). We next removed duplicated reads using Picard MarkDuplicate (1.141) for elimination of technical bias. iPSC and NPC chromatin accessible peaks were determined by MACS2 with specific option (macs2 callpeak --verbose 3 --treatment sample.bam -g hs -B -q 0.05 --extsize 200 --nomodel --shift -100 --nolambda --keep-dup all -f BAMPE --outdir dir --call-summits). To identify the confident stage-specific ATAC peaks, we employed a strategy that confirming background signal of each iPSC and NPC ATAC peaks. We first listed stage-specific peaks based macs2 peak calling results by selecting peaks which was not defined in other cell types. To define cell type-specific accessible regions, we further manipulated macs2 peak calling results. For each stage-specific peak, we plotted normalized ATAC peak intensity (log$_2$RPKM) distribution graph of corresponding cell types and another cell types simultaneously. We observed crossing point of two distribution plot and defined this point as background signal. Finally, we defined revised stage-specific ATAC peaks. For data visualization on genome browser, we used bamCoverage tool with following options; --normalizeUsing CPM --effectiveGenomeSize 2864785220.

Histone ChIP-seq library preparation
We executed H3K27ac ChIP-seq for profiling enhancers in iPSC and NPC cells. Cells were washed with Dulbecco's PBS (DPBS) and crosslinked by crosslinking buffer (100 mM NaCl, 0.1 mM EDTA, 5 mM HEPES [pH 8.0], 1% formaldehyde) for 10 min at room temperature. The crosslinking process was halted with 125 mM glycine in room temperature for 5 min with rotation and washed twice with cold DPBS. Crosslinked samples were resuspended in SDS lysis buffer (1% SDS, 50 mM Tris-HCl [pH 8.0], 10 mM EDTA) with protease inhibitor (Roche, Switzerland). After proceeding sonication (S220; Covaris, USA), the sonicated chromatin samples were incubated with H3K27ac antibody (39133; Active Motif, USA) and magnetic bead (10001D; Thermo Fisher Scientific, USA) for 4 h in 4°C with rotation. At the same time, input chromatin was stored from sonicated samples to be used as input control. After incubation, the bead complex was washed by RIPA-α (140 mM NaCl, 1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate, 10 mM Tris pH 8) with 4 times. The chromatin immunoprecipitated complex was treated with RNaseA (19101; Qiagen) and executed overnight reverse-crosslinking at 68°C. The chromatin immunoprecipitated DNA was extracted using

AMPure XP beads (A63881; Beckman Coulter), and libraries were prepared using NEB Next Ultra II DNA library Prep Kit (E7645; NEB, USA) following the manufacturer's instruction. The ChIP-seq libraries were sequenced in 150 bp paired-end mode using Illumina NovaSeq 6000 platform.

Histone ChIP-seq data processing
We utilized H3K4me3 and H3K27me3 ChIP-seq data for iPSC and NPC in Gene Expression Omnibus (GEO) (GSE156723). For read alignment, we used BWA (0.7.7) to map reads to the human genome sequence (hg19). We further removed duplicated reads using Picard MarkDuplicate (1.141). We further inspected the quality of the ChIP-seq experiment by measuring FRiP (fraction of reads in peaks) using ChIPQC (3.3). For peak calling, we used MACS2 (2.1.0) to call the significantly enriched ChIP region compare to the input IgG control "q -0.01 (FDR < 0.01%)". For the ATAC peak intensity comparison, we generated the normalized RPKM values by counting the number of reads for 50 bp bins then dividing by the total sequencing depth.

Motif analysis
To identify specifically enriched motifs in ATAC peak groups, we used the HOMER algorithm with default options (Heinz et al., 2010). We analyzed the known motifs results derived from published data which are provided by HOMER algorithm.

RNA-sequencing data processing
We utilized RNA-seq data for iPSC and NPC in GEO (GSE156723). Sequenced reads were mapped and aligned to the hg19 reference using Tophat2 (v2.1.0) and STAR (v2.6.0a) with the support of GRCh37 reference. From the mapped reads, we executed calculation of TPM (transcripts per million) using RSEM (v1.2.31) at the gene level. For data visualization browser, bamCoverage (deepTools 3.5.0) was used with default options.

In situ Hi-C library preparation
In situ Hi-C was performed on with iPSC and NPC cells, each comprising two biological replicates. About 1 million cells were used for crosslinking with 1% formaldehyde. *In situ* Hi-C was then conducted on the samples using protocol modified from previously study (Rao et al., 2014). Briefly, crosslinked cells were lysed (10 nM Tris-HCl [pH 8.0], 10 mM NaCl, and 0.2% IGEPAL CA630 [18896; Sigma-Aldrich]) and digested with 100U MboI (R0147; NEB). Digested fragments were labelled with biotin-14-dCTP (19518018; Invitrogen) and proximally ligated with T4 DNA Ligase (M0202; NEB), followed by reverse-crosslinking (Proteinase K [P8102; NEB] and 10% SDS). The fragments were then subjected to sonication (S220; Covaris) and purified with AMPure XP beads (A63880; Beckman Coulter). The final ligated DNA was pulled down with Dynabeads MyOne streptavidin T1 beads (65602; Invitrogen), followed by manual library preparation. Hi-C library was sequenced in 150 bp paired-end mode with Illumina HiSeq4000.

Hi-C data processing

Hi-C data processing is executed by described method in Yang et al. (2018). Raw reads were aligned to hg19 human reference genome. To focus on cis-chromosomal interactions, invalid Hi-C reads were removed by criteria which defined in previous studies; inter-chromosomal interactions, putative self-ligations, and non-restriction site specific ligations (Dixon et al., 2015; Schmitt et al., 2016). Valid cis-interactions were collected to build raw interaction frequency matrix with 5 kb resolution. A negative binomial model-based normalization pipeline removed biases of the interaction frequency matrix. Finally, distance dependent background signal was used to normalization for emphasizing biologically significant long-range chromatin interactions.

Data deposition

The NGS sequence data generated in this paper have been deposited in the GEO database, www.ncbi.nlm.nih.gov/geo (accession No. GSE158382).

RESULTS

Identification of chromatin accessibility alteration regions during iPSC to NPC differentiation

Before producing ATAC-seq data for iPSC to NPC, we confirmed our differentiation system by detecting pluripotency markers for iPSC (Oct4 and E-cadherin) and neuronal markers for NPC (Sox2 and Nestin) (Fig. 1A). To identify DNA regions that become accessible during neuronal differentiation, we performed ATAC-seq on both iPSCs and NPCs. We made biological replicates of the samples for ATAC-seq, confirmed their reproducibility, and proceeded to integrate them for further analysis (Fig. 1B). At first, we defined ATAC peaks in each cell type using MACS2 with specific options (Zhang et al., 2008). By using ChIPseeker with default settings (Yu et al., 2015), we observed that most of the ATAC peaks were found in the promoters and distal intergenic regions (Fig. 1C). Next, we categorized ATAC peaks into three groups depending on the existence of peaks in iPSC and NPC: common to both samples, iPSC-specific, or NPC-specific. We executed GREAT analysis for each ATAC peak groups with whole genome background (McLean et al., 2010) and checked the genomic positions of each group; 48.1% of Group C were observed within 5 kb from the transcription start sites (TSS), but in contrast, only 7.9% of Group I and 16.3% of Group N were located near the TSS regions (Fig. 1D). Interestingly, about 93% of Group I and about 83% of Group N were located at a distance over 5 kb from the TSS, suggesting that cell type-specific chromatin accessible regions may contain distal-acting regulatory elements (Fig. 1D). For ATAC peak grouping, we first used Deseq2 for differential peak calling. However, we found that the results could not reflect closed chromatin states (data not shown). So, we utilized the strategy of re-calling peaks from a previous study (Li et al., 2017). When we examined the RPKM (Reads Per Kilobase per Million mapped reads) of iPSC and NPC states in iPSC-specific peaks, we found that there was a clear separation point (1.75) between them. Thus, we set the peak intensity of 1.75 as our standard for the background signal (Fig. 1E). The

NPC-specific peaks also showed the same crossing point as the iPSC-specific peaks. Based on this standard, we refined the following three groups: commonly accessible regions (ATAC peak intensity > 1.75 in both iPSC and NPC: Group C), iPSC-specific accessible regions (ATAC peak intensity > 1.75 in iPSC and < in NPC: Group I), and NPC-specific accessible regions (ATAC peak intensity > 1.75 in NPC and < in iPSC: Group N).

Gene association with cell type-specific accessible chromatin regions

We next sought to determine which genes were linked to our defined groups. We performed hierarchical clustering analysis for the three ATAC peak groups and then examined which genes existed within +/- 5 kb from the ATAC peak regions (Fig. 2A). The genes in Group C were mainly house-keeping genes such as *ACTB*, *GAPDH*, *PGK1*, and *PPIA*. The genes in Group I were related to stemness including *POU5F1*, *ETS1*, and *KLF10*, while Group N included *HES1*, *NEUROG1*, *NEUROG2*, *PAX3*, and *PAX5*, which have roles in neuronal development and neural function. Using RNA-seq data, we confirmed that the gene expression was related to changes in chromatin accessibility (Fig. 2B). For example, the expression of ACTB and GAPDH from Group C were maintained in both iPSC and NPC samples as validated by RNA-seq analysis. Chromatin accessibility was high in the promoters of these genes in both iPSC and NPC samples (top of Fig. 2B). In the case of POU5F1 and ETS1, their expression level decreased after differentiation into NPC, and promoter chromatin accessibility consequently decreased (middle of Fig. 2B). By contrast, both expression level and promoter accessibility of NEUROG2 and HES1 increased after differentiation into NPC (bottom of Fig. 2B). We also performed gene ontology analysis with each group by DAVID (Fig. 2C). The genes of Group C were involved with cell division, cell cycle phase, and translation initiation, which are related to general cellular function. The genes of Group I were involved with somatic stem cell population maintenance. The genes from Group N were involved with neuronal differentiation and function terms such as midbrain development, Wnt signaling pathway, and axon guidance. Taken together, we concluded that our three ATAC peak groups correctly reflect the properties of each cell type.

Chromatin accessibility changes in enhancer reflected during iPSC to NPC differentiation

To further examine the features of each ATAC peak group, we decided to look into any histone modification changes corresponding to DNA accessibility during differentiation into NPC. To do this, we utilized three histone ChIP-seq data sets: H3K4me3 (active histone marker), H3K27me3 (repressive histone marker), and H3K27ac (active histone and enhancer marker). Group I showed a high level of H3K27me3 and low levels of H3K4me3 and H3K27ac compared with Group C (Fig. 3A). By contrast, Group N showed a lower level of H3K27me3 and elevated levels of H3K4me3 and H3K27ac, suggesting that changes in chromatin accessibility were positively linked to levels of H3K4me3 and H3K27ac and negatively to H3K27me3. For further analysis, we sub-categorized the groups into either promoter or enhancer. The promoter
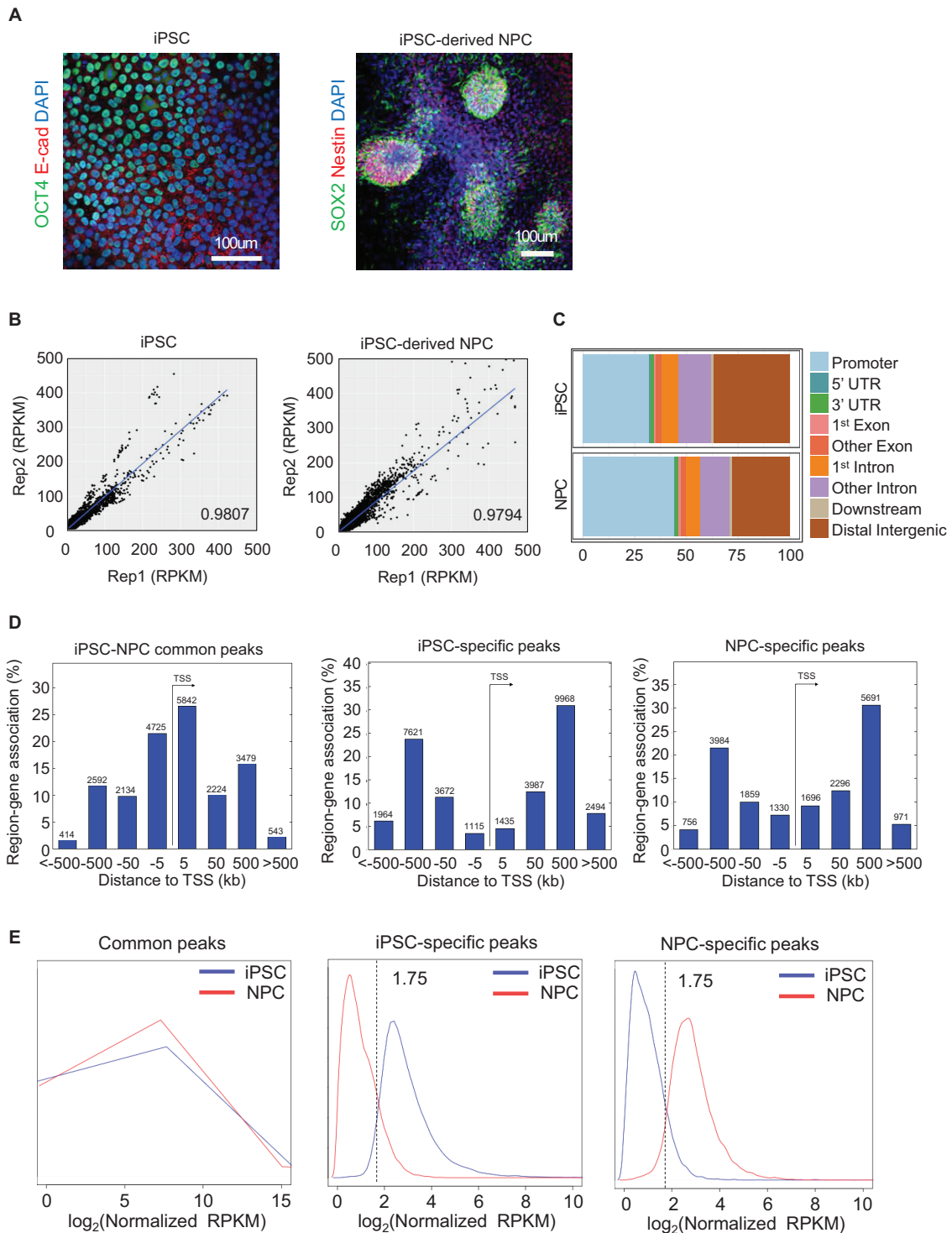
Fig. 1. Defining the iPSC and NPC-specific chromatin accessible regions by ATAC-seq. (A) Immunocytochemistry data of iPSC (OCT4 and E-Cadherin) and NPC (Nestin and Sox2). Scale bars = 100 μm. (B) Scatter plot of RPKM value for biological replicates of iPSC (left) and NPC (right) ATAC library. For all peaks defined by MACS2, RPKM of each cell type was calculated and compared with each other. Then we checked Pearson correlation coefficient. (C) Genomic distribution of iPSC and NPC ATAC peaks. The data was generated by 'ChIPseeker'. (D) Genomic localization of common (left), iPSC-specific (middle) and NPC-specific (right) ATAC peaks. The center of plots indicates transcription start site (TSS). (E) Density plots of the RPKM values for iPSC (blue) and NPC (red) samples. Three graphs are presented representing ATAC-seq common (left), iPSC-specific (middle), and NPC-specific (right) ATAC peaks. RPKM was normalized using a quantile normalization method. The crossing point of iPSC and NPC plots in the iPSC-specific and NPC-specific ATAC peaks is marked with a dotted line.
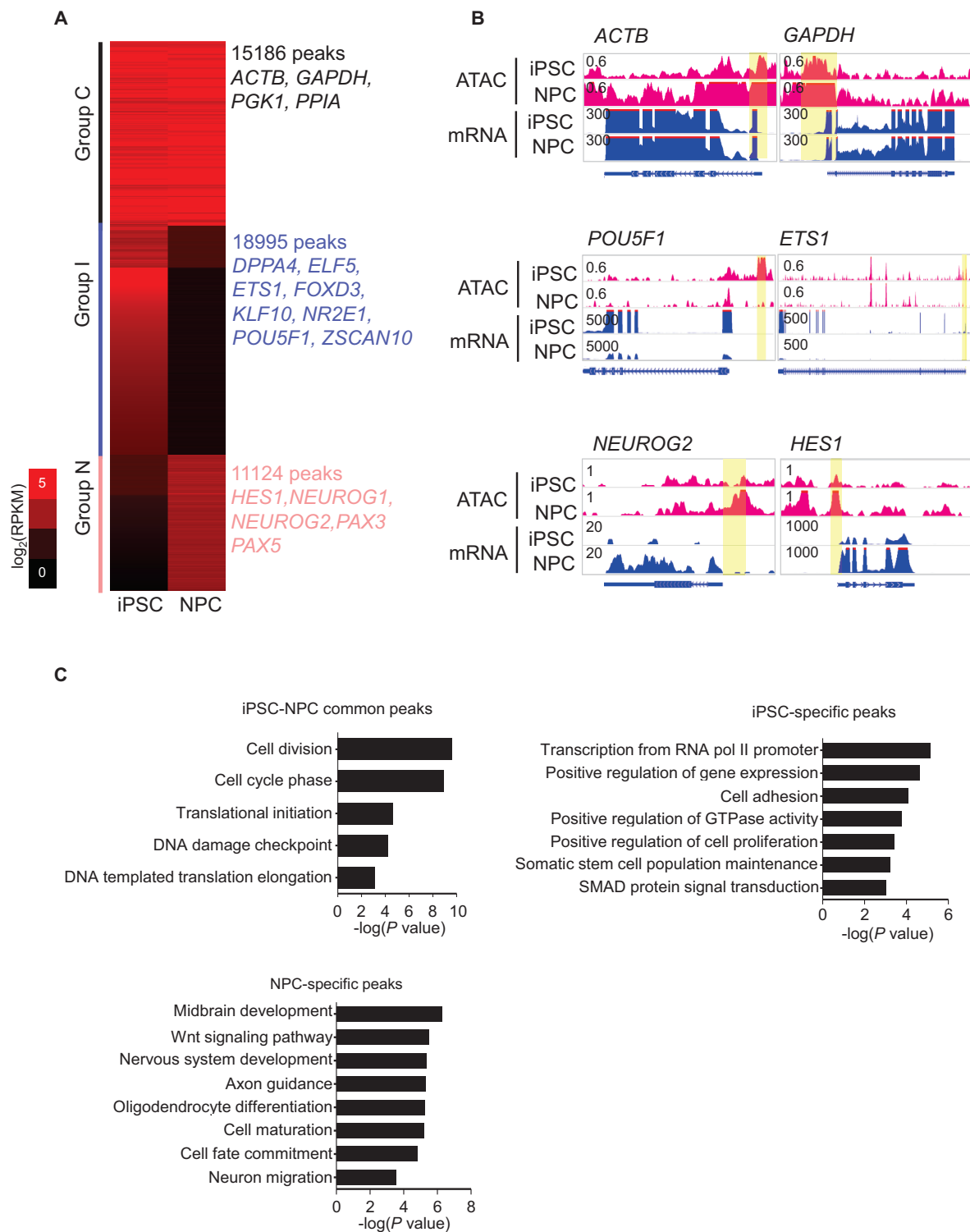
A

Group C — 15186 peaks
ACTB, GAPDH, PGK1, PPIA

Group I — 18995 peaks
DPPA4, ELF5, ETS1, FOXD3, KLF10, NR2E1, POU5F1, ZSCAN10

Group N — 11124 peaks
HES1, NEUROG1, NEUROG2, PAX3, PAX5

log₂(RPKM) 0–5

iPSC NPC

B

C

Fig. 2. Cell type-specific ATAC peaks and associated genes. (A) A heatmap of hierarchical clustering for iPSC and NPC ATAC-seq signals for common (Group C) and re-defined iPSC (Group I) and NPC (Group N) ATAC peaks based on Fig. 1. Group C involved house-keeping genes, Group I involved stem cell maintenance related genes, and Group N involved neural-related genes. (B) A graphical comparison of ATAC-seq and mRNA-seq reads for representative genes associated with Group C, Group I, and Group N is presented by the WashU Genome browser. The intensity of ATAC-seq and mRNA-seq read counts was calculated by command of bamCoverage in deepTools. (C) Results of gene ontology for common (upper left), iPSC-specific (upper right), and NPC-specific (lower left) ATAC peaks associated genes. Genes located in +/- 5 kb from ATAC peak regions were associated with each ATAC peak. The *P* value was calculated using the Benjamini-Hochberg method in DAVID.
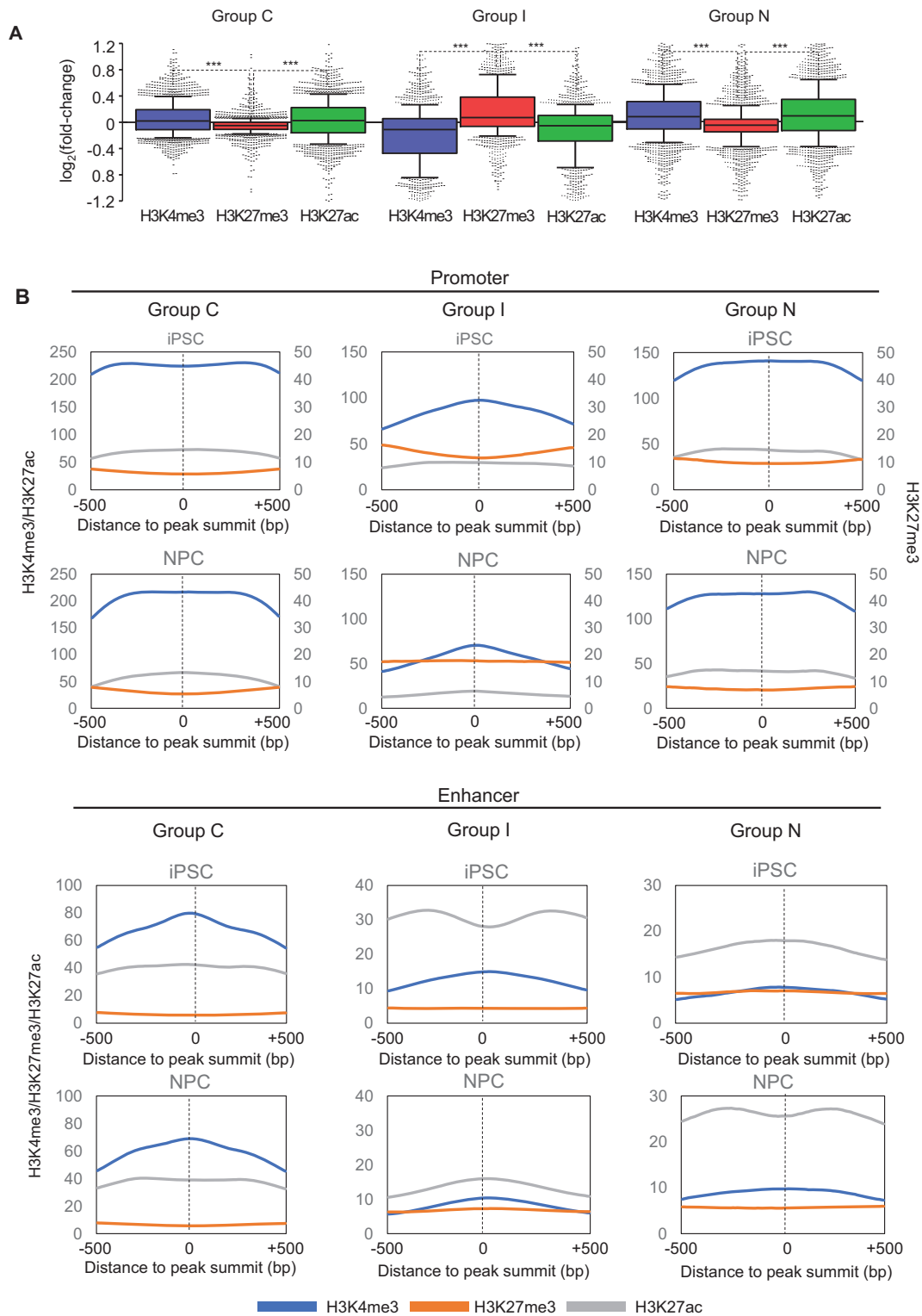
Fig. 3. Epigenetic modification of promoters and enhancer regions according to cell type-specific ATAC peaks. (A) Boxplots representing the differential level of H3K4me3, H3K27me3, and H3K27ac during iPSC to NPC differentiation in the common, iPSC-specific, and NPC-specific ATAC peak groups. Scale is $\log_2$(fold change). For the box-and-whisker plots, the upper and lower bounds of the boxes represent the 75th and 25th percentiles, respectively, and the horizontal lines indicate the median value. The upper and lower error bars indicate the 90th and 10th percentiles, respectively. ∗∗∗$P < 0.001$ (unpaired t-test, Mann-Whitney test, two-tailed). (B) Averaged RPKM signal intensity of H3K4me3, H3K27me3, and H3K27ac in common and cell type-specific ATAC peak groups. The graphs represent the histone modifications to promoters (upper) and enhancers (lower) of each ATAC peak group.

group was determined as the regions located within +/- 2 kb from the TSS, and the enhancer group was defined from the H3K27ac peak data with ENCODE candidate Regulatory Elements (cRE) data. We calculated the average levels of H3K4me3, H3K27me3, and H3K27ac respectively, and then profiled them in the regions of +/- 500 bp from the ATAC peak summit (Fig. 3B). During iPSC to NPC differentiation, the level of H3K4me3 in the promoter was decreased and the level of H3K27me3 was increased in the iPSC-specific peak (Group I), as expected. Surprisingly, no change was observed for H3K4me3 and H3K27me3 in Group N, even after differentiation into NPC. However, in the enhancer, the level of H3K27ac was relatively low in the iPSC-specific peak (Group I) and increased in the NPC-specific peak (Group N), corresponding with the differentiation process. Taken together, we concluded that the changes to chromatin accessibility have more effect on the activity of the enhancers rather than promoters.

NEUROD1 binds to the accessible regions of the enhancers during differentiation into NPC

It was reported that the cellular differentiation processes are regulated by specific TFs that initiate chromatin structure changes via cell type specific enhancers (Lupien et al., 2008). We performed TF motif enrichment analysis targeting our cell type specific ATAC peaks in the enhancer regions by HOMER analysis (Heinz et al., 2010). At first, we sought to find TFs that satisfy our criteria of P value ($< 10^{-10}$) and enrichment ratio (1.4) in each ATAC peak group. We clustered the TFs with -log (P value) and then found three groups: Group TF R (recurrence), Group TF I (iPSC-specific), and Group TF N (NPC-specific) (Fig. 4A). We found that the binding motifs for POU5F1 (OCT4) and SOX family members such as SOX2 and SOX9 were enriched in Group TF R. Although OCT4 and
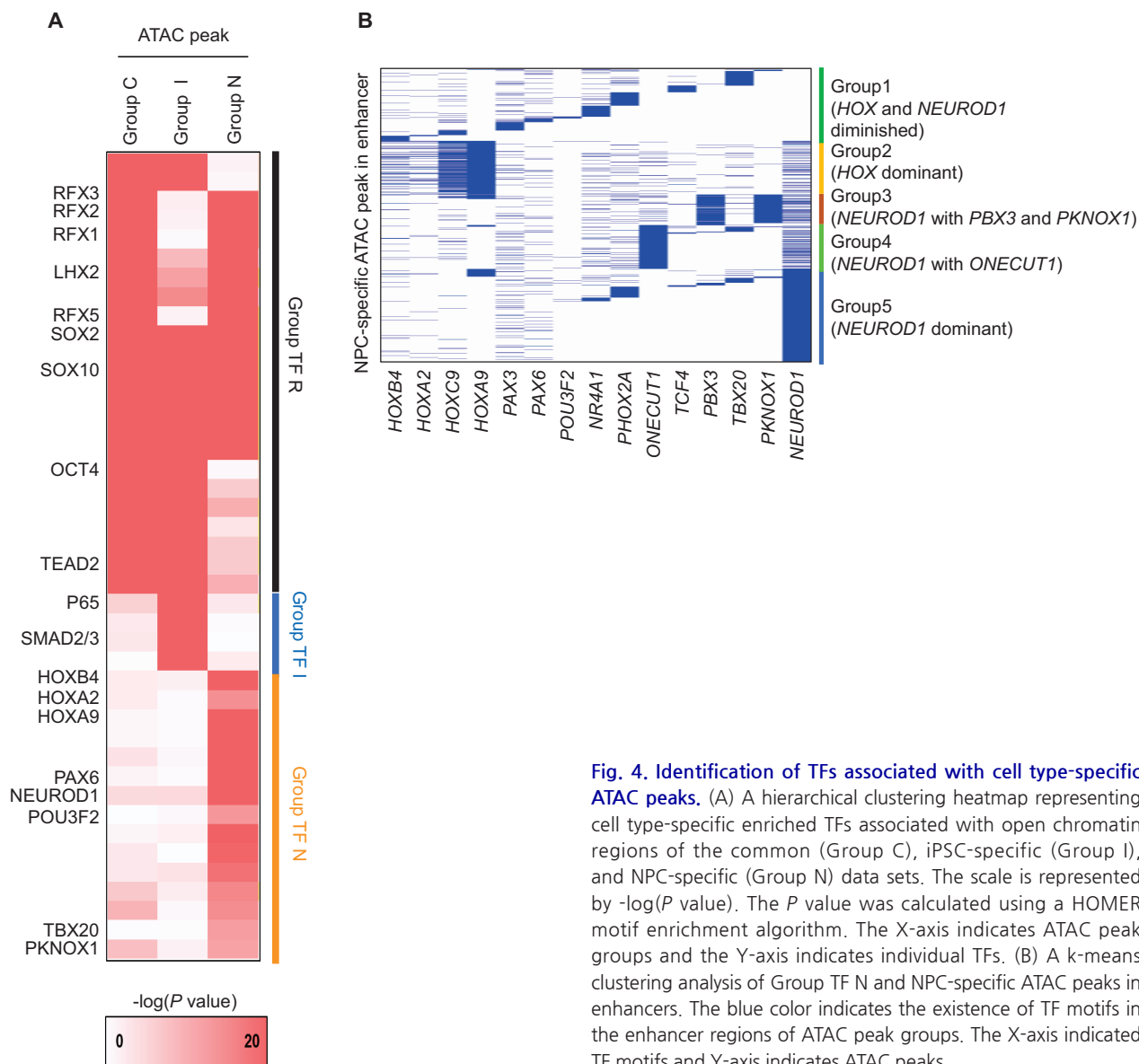

Fig. 4. Identification of TFs associated with cell type-specific ATAC peaks. (A) A hierarchical clustering heatmap representing cell type-specific enriched TFs associated with open chromatin regions of the common (Group C), iPSC-specific (Group I), and NPC-specific (Group N) data sets. The scale is represented by -log(P value). The P value was calculated using a HOMER motif enrichment algorithm. The X-axis indicates ATAC peak groups and the Y-axis indicates individual TFs. (B) A k-means clustering analysis of Group TF N and NPC-specific ATAC peaks in enhancers. The blue color indicates the existence of TF motifs in the enhancer regions of ATAC peak groups. The X-axis indicated TF motifs and Y-axis indicates ATAC peaks.

SOX2 are best known to induce pluripotency and maintain stemness, some papers also suggest crucial roles in NPC development (Cimadamore et al., 2011; Episkopou, 2005; Mitchell et al., 2014). Thus, it was not surprising that OCT4 and the SOX family were included in Group TF R. However, many homeo-domain containing genes (HOXA2, HOXA9, HOXB, HOXC9, PHOX2A, PBX3, PKNOX1, POU3F2 [Brn2]) and NEUROD1 were included in Group TF N, which suggests that their binding motifs were exclusively enriched in NPC-specific accessible DNA regions and rarely found in any other DNA regions. Zinc finger domain genes (NR4A1, ONE-CUT1) and paired box genes (PAX3, PAX6) were also found in Group TF N. To further understand the TFs in Group TF N, we performed a k-means clustering analysis. The associated TFs could be further classified into five groups: Group 1 (HOX and NEUROD1 diminished), Group 2 (HOX dominant), Group 3 (NEUROD1 with PBX1 and PKNOX1), Group 4 (NEUROD1 and ONECUT1), and Group 5 (NEUROD1 dominant) (Fig. 4B). The five groups were mainly classified by the degree of NEUROD1 expression; therefore, we concluded that NEUROD1 was a major determinant of NPC differentiation.

First line of genes that are induced during differentiation from iPSC into NPC

To characterize these five groups of TFs, we sought to determine which genes were connected to them by utilizing Hi-C data to analyze the physical chromatin interactions. We made an intra-chromosomal interaction matrix with 5 kb resolution, linked enhancers to target genes, and performed gene ontology analysis (Table 1). As a result, we found that Group 1 (HOX and NEUROD1 diminished) included genes related to organ morphogenesis, cellular differentiation, and the Hippo

signaling pathway. Group 2 (HOX dominant) included genes linked to neural tube development and the Wnt signaling pathway, and Group 3 (NEUROD1 with PBX1 and PKNOX1) were linked to the canonical Wnt signaling pathway and cell cycle related genes. Group 4 (NEUROD1 and ONECUT1) included genes related to neural crest cell migration, neural tube closure, and axon guidance. Lastly, Group 5 (NEUROD1 dominant) were linked to genes that promote cell proliferation, except keratinocytes, and promote the Notch signaling pathway (Table 1). Interestingly, Group 2 and Group 4 were related to neuronal development terms; however, overall the genes from Groups 1 to 5 were required for intrinsic NPC development from iPSCs.

DISCUSSION

Profiling chromatin accessibility changes during differentiation provides an insight into the regulatory mechanism of lineage-specific genes. These chromatin accessibility changes are mediated by a certain combination of TFs, so-called pioneer factors (Frank et al., 2015; Schaffner, 2015; Scott-Browne et al., 2016). The Yamanaka factors, which consist of OCT4, SOX2, NANOG, and MYC, are the most well-recognized pioneer factors that convert fibroblasts into stem cells (Takahashi and Yamanaka, 2006). Pioneer factors involved in several lineages have been discovered as well, including GATA4, MEF2C, and TBX5 which are able to induce direct reprogramming of fibroblasts into cardiomyocyte-like cells (Qian et al., 2013). Pioneer factors related to neuronal differentiation are also termed proneural factors, which mainly have bHLH structures (Baker and Brown, 2018; Bertrand et al., 2002). It is known that proneural factors, such as the BAM factors, NEUROD1, and NEUROG2, are able to bind nu-

Table 1. Target genes linked with enhancers that are associated with NPC-specific ATAC peaks

	Enhancer-TF group	P value	Genes
Group 1	NEUROD1 and HOX diminished		
	Organ morphogenesis	1.80E-02	GATA2, LHX3, SIX6 CDX1
	Cell differentiation	1.50E-02	DHCR7, DDX41, CSF1
	Hippo signaling pathway	1.90E-03	SOX2
Group 2	HOX dominant		
	Neural tube development	1.60E-03	EPHA2, TCF7
	Wnt signaling	2.20E-02	SKP1, CUL1, HHEX, MCC, SFRP2
Group 3	NEUROD1 with PBX3 and PKNOX1		
	Canonical Wnt signaling pathway	2.90E-02	NKX2-5, WNT7A
	Cell cycle	3.30E-02	PIM1, APEX2, CCND1
Group 4	NEUROD1 with ONECUT1		
	Neural crest cell migration	5.80E-04	EFNB1, GBX2
	Neural tube closure	9.10E-03	CECR2, LHX2, FZD3, SKI
	Axon guidance	3.30E-02	WNT3A, DVL1, GDF7
Group 5	NEUROD1 dominant		
	Positive regulation of cell proliferation	2.00E-02	CTBP2, MFNG, DTX4
	Negative regulation of keratinocyte proliferation	3.40E-02	EFNB2, EPPK1, KDF1, SFN
	Positive regulation of Notch signaling pathway	2.70E-02	GSX2, SLC35C2

We linked enhancers and genes by Hi-C interaction analysis with 5 kb resolution. Interactions with a ≥ 2-fold change above the background signal and a > 5 of normalized interaction counts were included in the analysis. P values were calculated using the Benjamini-Hochberg method in DAVID.

cleosomal chromatin and induce chromatin opening in fibroblasts during differentiation into neuronal cells (Soufi et al., 2012; Wapinski et al., 2013). The degree to which such TFs are able to bind to chromatin is reflected in chromatin accessibility, which then extrapolates to provide researchers with snapshots of dynamic regulatory capacity of the genome.

We used ATAC-seq, a widely adopted method given its robustness and relative ease (Klemm et al., 2019) to compare chromatin accessbility between iPSCs and its differentiated counterpart NPCs. Through analysis of chromatin status and histone modification data (H3K4me3, H3K27me3, H3K27ac), we showed that there is higher correlation between chromatin accessibility and activation of enhancers rather than with promoters. We then observed that the NEUROD1 binding sites were highly enriched in NPC-specific accessible regions in enhancers. In addition to NEUROD1, other TF motifs were identified such as the HOX family, PBX3, PHOX2, and ONECUT1, which are all involved in neuronal development. Generally, the HOX genes have conserved roles in body segmentation with dynamic expression patterns in response to cellular environment (Krumlauf, 1994; Philippidou and Dasen, 2013). The HOX 1 to 5 genes are highly expressed in the hindbrain, while HOX 4 to 11 are predominantly expressed in the spinal cord (Philippidou and Dasen, 2013). PHOX2 is known to play a role in neuronal development (Brunet and Pattyn, 2002; Dubreuil et al., 2002). PBX and PKNOX1 controls the localization each other, and PBX3 is highly expressed in the nervous system development (Ferretti et al., 2006; Rhee et al., 2004). In addition to the homeobox genes, the zinc finger proteins also play crucial roles in neuronal development (Lamar et al., 2001; Qin et al., 2009; Shimizu et al., 2010). It is known that NR4A1, a member of the nuclear receptor family of TFs, has a role in regulating neuronal development in the central nervous system, and together with ONECUT1/2, is highly expressed in the developing mouse retina (Akiyama et al., 2008; Honkaniemi and Sharp, 1999; Wu et al., 2012). Furthermore, PAX3 and PAX6 which belong to the paired box gene family, have distinctive roles in early neuronal differentiation and the developing spinal cord (Bel-Vialar et al., 2007; Terzic and Saraga-Babic, 1999). Finally, based on physical chromatin interactions with the above TFs, we were able to identify the first line of genes involved in neuronal lineage specification. Our findings carry significance—while various studies involving iPSC-derived neuronal disease models have suggested causal relationship between non-coding disease variants and defect of regulatory elements, giving rise to dysregulated TF binding to neuronal cell-type specific accessible regions in neuronal diseases such as Alzheimer's disease and schizophrenia (Forrest et al., 2017; Schwartzentruber et al., 2018; Zhang et al., 2020)—there still remains a lack of information on the downstream genes that are affected by defective regulatory elements. Cell fate specification consists of multiple steps, necessitating precise control of gene regulation (Frank et al., 2015) and we have provided a framework that is able to accurately identify genes involved in early neuronal differentiation from iPSCs to NPCs through the integrated analysis of chromatin accessibility, histone modification and chromatin spatial organization.

While the role of NEUROD1 in inducing changes to chro-matin accessibility during direct reprogramming process (Matsuda et al., 2019) has been recognized, other factors that cooperate with NEUROD1 are still unknown, particularly during the intrinsic developmental process from iPSC into NPC. To the best of our knowedege, our study is the first report to fully lineate the following: 1) localization of chromatin sites that are first opened, 2) corresponding TFs that bind to these sites, and 3) initial genes under regulation of the TFs during neuronal differentiation. Although there remains a need to characterize expression dynamics of NeuroD1 during iPSC to NPC differentiation, this study sheds light on the understanding of normal early neuronal development and holds potential to be used as a resource in comparing with neuronal disease model using iPSC.

AUTHOR CONTRIBUTIONS
W.Y.C. and Y.J.K. conceived and designed the study. W.Y.C. and A.J.L. produced and processed next generation sequencing data and W.Y.C. analyzed the data. A.N.C. performed cellular work for iPSC maintaining and differentiation. I.J. and S.W.C. contributed advice in experiment and data analysis. W.Y.C., J.H.H., and L.K.K. wrote the manuscript. L.K.K. and Y.J.K. supervised the project.

CONFLICT OF INTEREST
The authors have no potential conflicts of interest to disclose.

ORCID
Won-Young Choi   https://orcid.org/0000-0002-3576-5682
Ji-Hyun Hwang   https://orcid.org/0000-0001-6009-0632
Ann-Na Cho   https://orcid.org/0000-0003-0047-8867
Andrew J. Lee   https://orcid.org/0000-0001-9917-8834
Inkyung Jung   https://orcid.org/0000-0002-5885-2754
Seung-Woo Cho   https://orcid.org/0000-0001-8058-332X
Lark Kyun Kim   https://orcid.org/0000-0001-5983-4470
Young-Joon Kim   https://orcid.org/0000-0001-5061-587X

REFERENCES

Akiyama, K., Ishikawa, M., and Saito, A. (2008). mRNA expression of activity-regulated cytoskeleton-associated protein (arc) in the amygdala-kindled rats. Brain Res. 1189, 236-246.

Baker, N.E. and Brown, N.L. (2018). All in the family: proneural bHLH genes and neuronal diversity. Development 145, dev159426.

Bel-Vialar, S., Medevielle, F., and Pituello, F. (2007). The on/off of Pax6 controls the tempo of neuronal differentiation in the developing spinal cord. Dev. Biol. 305, 659-673.

Bertrand, N., Castro, D.S., and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. Nat. Rev. Neurosci. 3, 517-530.

Brunet, J.F. and Pattyn, A. (2002). Phox2 genes - from patterning to connectivity. Curr. Opin. Genet. Dev. 12, 435-440.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr. Protoc. Mol. Biol. 109, 21.29.1-21.29.9.

Cedar, H. and Bergman, Y. (2009). Linking DNA methylation and histone

modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295-304.

Chang, C.Y., Ting, H.C., Liu, C.A., Su, H.L., Chiou, T.W., Lin, S.Z., Harn, H.J., and Ho, T.J. (2020). Induced pluripotent stem cell (iPSC)-based neurodegenerative disease models for phenotype recapitation and drug screening. Molecules *25*, 2000.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. Nat. Methods *8*, 424-429.

Chen, Y.M., Chen, L.H., Li, M.P., Li, H.F., Higuchi, A., Kumar, S.S., Ling, Q.D., Alarfaj, A.A., Munusamy, M.A., Chang, Y., et al. (2017). Xeno-free culture of human pluripotent stem cells on oligopeptide-grafted hydrogels with various molecular designs. Sci. Rep. *7*, 45146.

Cimadamore, F., Fishwick, K., Giusto, E., Gnedeva, K., Cattarossi, G., Miller, A., Pluchino, S., Brill, L.M., Bronner-Fraser, M., and Terskikh, A.V. (2011). Human ESC-derived neural crest model reveals a key role for SOX2 in sensory neurogenesis. Cell Stem Cell *8*, 538-551.

Cotovio, J.P. and Fernandes, T.G. (2020). Production of human pluripotent stem cell-derived hepatic cell lineages and liver organoids: current status and potential applications. Bioengineering (Basel) *7*, 36.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dubreuil, W., Hirsch, M.R., Jouve, C., Brunet, J.F., and Goridis, C. (2002). The role of Phox2b in synchronizing pan-neuronal and type-specific aspects of neurogenesis. Development *129*, 5241-5253.

Episkopou, V. (2005). SOX2 functions in adult neural stem cells. Trends Neurosci. *28*, 219-221.

Ferretti, E., Villaescusa, J.C., Di Rosa, P., Fernandez-Diaz, L.C., Longobardi, E., Mazzieri, R., Miccio, A., Micali, N., Selleri, L., Ferrari, G., et al. (2006). Hypomorphic mutation of the TALE gene Prep1 (pKnox1) causes a major reduction of Pbx and Meis proteins and a pleiotropic embryonic phenotype. Mol. Cell. Biol. *26*, 5650-5662.

Forrest, M.P., Zhang, H., Moy, W., McGowan, H., Leites, C., Dionisio, L.E., Xu, Z., Shi, J., Sanders, A.R., Greenleaf, W.J., et al. (2017). Open chromatin profiling in hiPSC-derived neurons prioritizes functional noncoding psychiatric risk variants and highlights neurodevelopmental loci. Cell Stem Cell *21*, 305-318.e8.

Frank, C.L., Liu, F., Wijayatunge, R., Song, L.Y., Biegler, M.T., Yang, M.G., Vockley, C.M., Safi, A., Gersbach, C.A., Crawford, G.E., et al. (2015). Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. Nat. Neurosci. *18*, 647-656.

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, A., Kao, S.H., Ling, Q.D., Chen, Y.M., Li, H.F., Alarfaj, A.A., Munusamy, M.A., Murugan, K., Chang, S.C., Lee, H.C., et al. (2015). Long-term xeno-free culture of human pluripotent stem cells on hydrogels with optimal elasticity. Sci. Rep. *5*, 18136.

Honkaniemi, J. and Sharp, F.R. (1999). Prolonged expression of zinc finger immediate-early gene mRNAs and decreased protein synthesis following kainic acid induced seizures. Eur. J. Neurosci. *11*, 10-17.

Jin, Y., Lee, J.U., Chung, E., Yang, K., Kim, J., Kim, J.W., Lee, J.S., Cho, A.N., Oh, T., Lee, J.H., et al. (2019). Magnetic Control of axon navigation in reprogrammed neurons. Nano Lett. *19*, 6517-6523.

Kang, S., Chen, X., Gong, S., Yu, P., Yau, S., Su, Z., Zhou, L., Yu, J., Pan, G., and Shi, L. (2017). Characteristic analyses of a neural differentiation model from iPSC-derived neuron according to morphology, physiology, and global gene expression pattern. Sci. Rep. *7*, 12233.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207-220.

Krumlauf, R. (1994). Hox genes in vertebrate development. Cell *78*, 191-201.

Lamar, E., Kintner, C., and Goulding, M. (2001). Identification of NKL, a novel Gli-Kruppel zinc-finger protein that promotes neuronal differentiation. Development *128*, 1335-1346.

Li, D., Liu, J., Yang, X., Zhou, C., Guo, J., Wu, C., Qin, Y., Guo, L., He, J., Yu, S., et al. (2017). Chromatin accessibility dynamics during iPSC reprogramming. Cell Stem Cell *21*, 819-833.e6.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958-970.

Matsuda, T., Irie, T., Katsurabayashi, S., Hayashi, Y., Nagai, T., Hamazaki, N., Adefuin, A.M.D., Miura, F., Ito, T., Kimura, H., et al. (2019). Pioneer factor NeuroD1 rearranges transcriptional and epigenetic profiles to execute microglia-neuron conversion. Neuron *101*, 472-485.e7.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495-501.

Mitchell, R.R., Szabo, E., Benoit, Y.D., Case, D.T., Mechael, R., Alamilla, J., Lee, J.H., Fiebig-Comyn, A., Gillespie, D.C., and Bhatia, M. (2014). Activation of neural cell fate programs toward direct conversion of adult human fibroblasts into tri-potent neural progenitors using OCT-4. Stem Cells Dev. *23*, 1937-1946.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C., et al. (2011). Induction of human neuronal cells by defined transcription factors. Nature *476*, 220-223.

Pataskar, A., Jung, J., Smialowski, P., Noack, F., Calegari, F., Straub, T., and Tiwari, V.K. (2016). NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. EMBO J. *35*, 24-45.

Philippidou, P. and Dasen, J.S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. Neuron *80*, 12-34.

Qian, L., Berry, E.C., Fu, J.D., Ieda, M., and Srivastava, D. (2013). Reprogramming of mouse fibroblasts into cardiomyocyte-like cells in vitro. Nat. Protoc. *8*, 1204-1215.

Qin, Z., Ren, F., Xu, X., Ren, Y., Li, H., Wang, Y., Zhai, Y., and Chang, Z. (2009). ZNF*536*, a novel zinc finger protein specifically expressed in the brain, negatively regulates neuron differentiation by repressing retinoic acid-induced gene transcription. Mol. Cell. Biol. *29*, 3633-3643.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rhee, J.W., Arata, A., Selleri, L., Jacobs, Y., Arata, S., Onimaru, H., and Cleary, M.L. (2004). Pbx3 deficiency results in central hypoventilation. Am. J. Pathol. *165*, 1343-1350.

Ross, S.E., Greenberg, M.E., and Stiles, C.D. (2003). Basic helix-loop-helix factors in cortical development. Neuron *39*, 13-25.

Rowe, R.G. and Daley, G.Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Genet. *20*, 377-388.

Rue, P. and Martinez Arias, A. (2015). Cell dynamics and gene expression control in tissue homeostasis and development. Mol. Syst. Biol. *11*, 792.

Schaffner, W. (2015). Enhancers, enhancers - from their discovery to today's universe of transcription enhancers. Biol. Chem. *396*, 311-327.

Schmitt, A.D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L., et al. (2016). A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep. *17*, 2042-2059.

Schwartzentruber, J., Foskolou, S., Kilpinen, H., Rodrigues, J., Alasoo, K., Knights, A.J., Patel, M., Goncalves, A., Ferreira, R., Benn, C.L., et al. (2018).

Molecular and functional variation in iPSC-derived sensory neurons. Nat. Genet. *50*, 54-61.

Scott-Browne, J.P., Lopez-Moyado, I.F., Trifari, S., Wong, V., Chavez, L., Rao, A., and Pereira, R.M. (2016). Dynamic changes in chromatin accessibility occur in CD8(+) T cells responding to viral infection. Immunity *45*, 1327-1340.

Seo, H.I., Cho, A.N., Jang, J., Kim, D.W., Cho, S.W., and Chung, B.G. (2015). Thermo-responsive polymeric nanoparticles for enhancing neuronal differentiation of human induced pluripotent stem cells. Nanomedicine *11*, 1861-1869.

Shimizu, T., Nakazawa, M., Kani, S., Bae, Y.K., Shimizu, T., Kageyama, R., and Hibi, M. (2010). Zinc finger genes Fezf1 and Fezf2 control neuronal differentiation by repressing Hes5 expression in the forebrain. Development *137*, 1875-1885.

Shin, J., Choi, E.J., Cho, J.H., Cho, A.N., Jin, Y., Yang, K., Song, C., and Cho, S.W. (2017). Three-dimensional electroconductive hyaluronic acid hydrogels incorporated with carbon nanotubes and polypyrrole by catechol-mediated dispersion enhance neurogenesis of human neural stem cells. Biomacromolecules *18*, 3060-3072.

Shum, C., Macedo, S.C., Warre-Cornish, K., Cocks, G., Price, J., and Srivastava, D.P. (2015). Utilizing induced pluripotent stem cells (iPSCs) to understand the actions of estrogens in human neurons. Horm. Behav. *74*, 228-242.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Terzic, J. and Saraga-Babic, M. (1999). Expression pattern of PAX3 and PAX6 genes during human embryogenesis. Int. J. Dev. Biol. *43*, 501-508.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., et al. (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621-635.

Wu, F., Sapkota, D., Li, R., and Mu, X. (2012). Onecut 1 and Onecut 2 are potential regulators of mouse retinal development. J. Comp. Neurol. *520*, 952-969.

Wu, Y.Y., Chiu, F.L., Yeh, C.S., and Kuo, H.C. (2019). Opportunities and challenges for the use of induced pluripotent stem cells in modelling neurodegenerative disease. Open Biol. *9*, 180177.

Yang, D., Jang, I., Choi, J., Kim, M.S., Lee, A.J., Kim, H., Eom, J., Kim, D., Jung, I., and Lee, B. (2018). 3DIV: a 3D-genome Interaction Viewer and database. Nucleic Acids Res. *46*, D52-D57.

Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383.

Zhang, S.W., Zhang, H.W., Zhou, Y.F., Qiao, M., Zhao, S.M., Kozlova, A., Shi, J.X., Sanders, A.R., Wang, G., Luo, K.X., et al. (2020). Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. Science *369*, 561-565.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785-798.