**ANALYSIS OF DNA CPG METHYLATION OF RETINAL DIFFERENTIATION FROM HUMAN BLASTOMERE-LIKE STEM CELLS**

by

Xiaoyuan Chen

A thesis submitted to Johns Hopkins University in conformity with the requirement for the degree of Master of Science in Engineering

Baltimore, Maryland

May 2025

© 2025 Xiaoyuan Chen

All rights reserved

# Abstract

Human pluripotent stem cells (hPSCs) promise new routes to regenerative therapies, yet the epigenetic rules that steer them toward specific fates are still hazy. Here we focus on a blastomere‑like stem cell state produced by tankyrase/PARP inhibition (TIRN‑SCs)—an even earlier stage than the classic naïve state—and track how it specifies the retinal lineage. Using multi-omics data combining RNA‑seq, whole‑genome bisulfite sequencing, and targeted ChIP‑seq, we map CpG methylation and other epigenetic changes through retinal organoid differentiation comparing TIRN and primed pluripotency state. Surprisingly, our results reveal that TIRN-SC do not exhibit global loss of methyl marks the way many naïve PSCs are reported to be. Instead, they undergo demethylation highly selectively: more than 10,000 differentially methylated regions (DMRs) identified to cluster in CpG shores, promoters, and other key regulatory elements while the rest of the genome stays almost unchanged. Those putative regions coincide with the differential binding sites of the pluripotency trio — SOX2, OCT4, and NANOG — hinting that local methyl loss may regulate the chromatin accessibility for these factors to reach and maintain in the blastomere‑like program.

Further analysis leads to a more comprehensive rationale combining out data with prior proteogenomic work suggesting that reduced PARP1 and elevated tankyrase levels appear to release SON factors from repressive complexes and to enable their redistribution to newly demethylated sites. Many of these sites are identified within or near enhancer elements that transcribe short non‑coding RNAs that are recognized as putative enhancer RNA/super-enhancer RNA (eRNA/seRNA) signatures based on enhanced H3K27ac signaling, presumably inducing eventual blastomere-like transcriptomics dynamics. Consistent with the prior work, these results imply an epigenetic and proteomic, integrated regulatory network where methylation loss, SON binding, and enhancer activation jointly amplify early developmental circuits. Despite these significant epigenomic remodeling of TIRN-SC, transcriptomics programs correlate poorly with CpG methylation status, suggesting again the multi‑layered indirect regulation. After retinal differentiation, most methylation and expression differences between TIRN‑SCs and primed cells faded, underscoring the complexity of TIRN pluripotency dynamics and the need to probe post‑transcriptional controls in retinal organoid development.

**Primary Reader:** Elias T Zambidis

**Committee Member:** Patrick Cahan

**Committee Member:** Michael J Betenbaugh

# Acknowledgements

I would like to express my deepest gratitude to those who have supported and guided me throughout the course of my thesis work.

First and foremost, I am profoundly grateful to **Dr. Elias T Zambidis**, my principal investigator, for welcoming me into his lab and giving me the opportunity to explore the fascinating field of naïve stem cells. His invaluable advice, insightful feedback, and unwavering support have been instrumental in the development and completion of this thesis.

I would also like to sincerely thank **Dr. Ludovic Zimmerlin** for his detailed and patient guidance on bioinformatics techniques. His generous help in connecting my project to the foundational work previously done in the lab has been crucial in shaping the direction of my research.

My heartfelt appreciation goes to **Dr. Michael J Betenbaugh**, who served as a member of my thesis committee and kindly allowed me to gain additional research experience in his lab. His support has broadened my scientific perspective and enriched my academic journey.

Lastly, I would like to thank **Dr. Patrick Cahan** for serving on my thesis committee and for teaching the Computational Stem Cell Biology course, which provided me with essential knowledge that greatly inspired smy work.

To all of you, thank you for your mentorship, encouragement, and generosity.

# Dedication

This thesis is dedicated to my parents, **Zhaolong Chen** and **Chunlan Wang**, for their constant encouragement, wholehearted support, and unreserved love.

Contents

[Abstract ii](#_Toc196179040)

[Acknowledgements iv](#_Toc196179041)

[Dedication v](#_Toc196179042)

[List of Tables viii](#_Toc196179043)

[List of Figures ix](#_Toc196179044)

[Introduction 1](#_Toc196179045)

[Methods 4](#_Toc196179046)

[Cell culture 4](#_Toc196179047)

[Undifferentiated cell culture 4](#_Toc196179048)

[Retinal organoid cultrue 5](#_Toc196179049)

[Bioinformatics 6](#_Toc196179050)

[Bulk RNA-Seq studies 6](#_Toc196179051)

[ChIP-Seq studies 8](#_Toc196179052)

[WGBS studies 10](#_Toc196179053)

[Results 15](#_Toc196179054)

[TIRN-SC exhibited a modest global hypomethylation that was reduced or slightly reversed upon differentiation into retinal organoids. 15](#_Toc196179055)

[Methylation differences were significantly expanded at differentially methylated regions (DMRs) in the undifferentiated state. 17](#_Toc196179056)

[Local demethylation at core transcription factor binding sites drove the blastomere-like pluripotency of TIRN-SC. 25](#_Toc196179057)

[The transcriptional programs of TIRN-SC were not directly regulated by CpG methylation. 29](#_Toc196179058)

[Core transcription factors mediated the CpG methylation regulation of transcription by activating putative enhancer and super-enhancer. 37](#_Toc196179059)

[Discussion 43](#_Toc196179060)

[Limitations 47](#_Toc196179061)

[Bibliography 48](#_Toc196179062)

# List of Tables

[**Table 1.** Sample design matrix used in ‘DMRichR’ package 12](#_Toc196171099)

[**Table 2.** Genome wide-CpG methylation levels of individual cell lines 16](#_Toc196171100)

[**Table 3.** ChromHMM 25-state model 42](#_Toc196171101)

# List of Figures

[**Figure 1.** Average genome wide-CpG methylation levels of TIRN and primed cell lines in the undifferentiated state and RO 16](#_Toc196171124)

[**Figure 2.** Genome wide-CpG methylation levels of individual cell lines 16](#_Toc196171125)

[**Figure 3.** Volcano plot showing percent difference of CpG methylation of DMRs between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state 19](#_Toc196171126)

[**Figure 4.** Smoothed methylation curve of sample DMR on exon and CpG shore between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state. 19](#_Toc196171127)

[**Figure 5.** Smoothed methylation curve of sample DMR on genic downstream and CpG shore between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state. 20](#_Toc196171128)

[**Figure 6.** Smoothed methylation curve of sample DMR on intergenic regions and open sea between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state. 20](#_Toc196171129)

[**Figure 7.** Enrichment analysis on CpG context of DMRs between TIRN-SC and primed hPSC in the undifferentiated state. 20](#_Toc196171130)

[**Figure 8.** Enrichment analysis on genic regions of DMRs between TIRN-SC and primed hPSC in the undifferentiated state. 21](#_Toc196171131)

[**Figure 9.** Volcano plot showing percent difference of CpG methylation of DMRs between TIRN-SC (red) and primed (blue) hPSC in RO 22](#_Toc196171132)

[**Figure 10**. Smoothed methylation curve of sample DMR on exon and CpG island between TIRN-SC (red) and primed (blue) hPSC in RO. 23](#_Toc196171133)

[**Figure 11.** Smoothed methylation curve of sample DMR on genic downstream and CpG island-shore combined between TIRN-SC (red) and primed (blue) hPSC in RO. 23](#_Toc196171134)

[**Figure 12.** Smoothed methylation curve of sample DMR on intergenic regions and open sea between TIRN-SC (red) and primed (blue) hPSC in RO. 24](#_Toc196171135)

[**Figure 13.** Enrichment analysis on CpG context of DMRs between TIRN-SC and primed hPSC in RO. 24](#_Toc196171136)

[**Figure 14.** Enrichment analysis on genic regions of DMRs between TIRN-SC and primed hPSC in RO. 25](#_Toc196171137)

[**Figure 15.** Profile plot of WGBS methylation scores on the differential NANOG binding, SOX2 binding, OCT4 binding, and SON co-binding regions between TIRN-SC and primed hPSC in the undifferentiated state 26](#_Toc196171138)

[**Figure 16.** Profile plot of WGBS methylation scores on 100,000 random genomic regions (400 bp) between TIRN-SC and primed hPSC in the undifferentiated state 27](#_Toc196171139)

[**Figure 17.** Profile plot of WGBS methylation scores on the differential NANOG binding, SOX2 binding, OCT4 binding, and SON co-binding regions between TIRN-SC and primed hPSC in RO 28](#_Toc196171140)

[**Figure 18.** Profile plot of WGBS methylation scores on 100,000 random genomic regions (400 bp) between TIRN-SC and primed hPSC in RO. 29](#_Toc196171141)

[**Figure 19.** Profile plot of WGBS methylation scores on the differential SON co-binding regions between TIRN-SC and primed hPSC during retinal differentiation 29](#_Toc196171142)

[**Figure 20.** Volcano plot showing log2 fold change of RNA expression (**left**) percent difference of CpG methylation of DMRs (**right**) between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state 30](#_Toc196171143)

[**Figure 21.** Heatmaps of RNA-seq data (FPKM) for TIRN-SC and primed hPSC lines comparing key transcription factors expressed during human embryonic stages from pre-ZGA to ICM 31](#_Toc196171144)

[**Figure 22.** Reactome GSEA pathway analysis of differential RNA expressions between TIRN and primed hPSC in the undifferentiated state. 32](#_Toc196171145)

[**Figure 23.** Cross plot of log2 fold change of RNA-Seq (x-axis; RNA) vs percent difference of WGBS (y-axis; methylation) between TIRN and primed hPSC in the undifferentiated state. 33](#_Toc196171146)

[Figure 24. Volcano plot showing log2 fold change of RNA expression (**left**) percent difference of CpG methylation of DMRs (**right**) between TIRN-SC (red) and primed (blue) hPSC in RO 34](#_Toc196171147)

[**Figure 25.** Heatmaps of RNA-seq data (FPKM) for TIRN-SC and primed hPSC lines comparing key transcription factors expressed in human retinal differentiation 35](#_Toc196171148)

[**Figure 26.** Reactome GSEA pathway analysis of differential RNA expressions between TIRN and primed hPSC in RO. 36](#_Toc196171149)

[**Figure 27.** Cross plot of log2 fold change of RNA-Seq (x-axis; RNA) vs percent difference of WGBS (y-axis; methylation) between TIRN and primed hPSC in RO. 37](#_Toc196171150)

[**Figure 28.**Sample IGV snapshots of RNA-Seq and ChIP-Seq score tracks for corresponding DMR in the undifferentiated state. 41](#_Toc196171151)

# Introduction

Human pluripotent stem cells (hPSCs), including human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs), are featured by their unique ability of unlimited self-renewal and differentiation potential into multiple lineages, providing irreplaceable models for regenerative medicine and developmental biology research1,2. These cells can be primarily divided into two pluripotent states: naïve and primed3. The naïve state is typically represented by cells from the pre-implantation blastocyst inner cell mass and maintains an earlier stage of development, while the primed state describes the pluripotency of the post-implantation epiblast cells with restricted developmental potential in the primed lineage4,5. However, recent research by Zimmerlin *et al* shows that tankyrase/PARP inhibitor-regulated naïve stem cells (TIRN-SCs) can be reverted to the developmentally even earlier blastomere-like stage, which provides a more potent differentiation capacity and a closer reach to totipotency6. Understanding the epigenetic mechanisms reaching and maintaining such enhanced pluripotency, especially DNA methylation, is thus necessary for optimizing differentiation protocols and further contributing to their enhanced therapeutic potential.

DNA methylation is the process of adding a methyl group predominantly covalently to cytosines within CpG dinucleotides by DNA methyltransferase (DNMT) and is a key regulation mechanism in multiple molecular programs, including gene expression, chromatin structure, and genomic stability7. Significant differences in global DNA methylation levels between naïve and primed hPSCs have been previously reported. Naïve hPSCs tend to experience significant global hypomethylation that is typically explained as the reason for their enhanced differentiation potential compared to the primed hPSCs8,9. However, increasing new evidence suggests that extensive demethylation in naïve hPSCs can cause unexpected consequences, such as loss of key methylation imprinting at specific regulatory regions, which cannot be retrieved after differentiation and therefore decreases differentiation quality and genomic stability10,11.

While global methylation changes have been well-documented, the local methylation differences at specific regulatory regions are less studied for their potential interactions with other spatially overlapped regulatory elements, such as transcription factors. Core transcription factors, including NANOG, SOX2, and OCT4, play essential roles in regulating the pluripotency dynamics based on their unique functions of remodeling chromatin structures, modulating gene expression, influencing protein homeostasis, and ultimately defining cellular identity12,13. Recent studies by Zimmerlin *et al* demonstrate that TIRN-SCs reprogramming with distinct genome-wide proteogenomic profiles are attributed to the interacted relationship between transcription factor binding, chromatin state, and DNA methylation6. Investigating these interactions at specific regulatory regions can provide novel understandings of how naïve pluripotency is maintained, enhanced, and effectively utilized during differentiation at molecular level.

Considering the promising therapeutic approach for degenerative eye diseases, such as age-related macular degeneration and retinitis pigmentosa14, the advancements in retinal differentiation from hPSCs can be highly valuable. Therefore, it is necessary to track the dynamic changes in the transcriptomics and epigenomics during such differentiation with a focus on key molecular differences that potentially lead to advantages of naïve hPSCs in terms of final tissue functions. Notably, recent work from Zambidis et al has suggested that TIRN-SCs can provide morphological and functional superiority of the retinal organoids (RO) developed from them and thus improve the generation efficiency and engraftment lifespan compared to primed cells15,16. Exploring the significant changes of DNA methylation patterns and its association with other regulatory element dynamics through the retinal differentiation from functional blastomere-like TIRN-SCs can therefore make important contributions to optimizing differentiation protocols and improving clinical outcomes.

In this study, I analyzed DNA CpG methylation change differences between naïve and primed hPSCs in retinal differentiation process using the functional human blastomere-like TIRN-SC as a novel platform. Combining RNA-seq, whole-genome bisulfite sequencing (WGBS), and chromatin immunoprecipitation sequencing (ChIP-seq), I conduct an integrated investigation of how global and locus-specific methylation dynamics correlate with transcription factor binding patterns and gene expression programs during differentiation. Contrary to previous findings of extensive global hypomethylation in naïve stem cells8,9, my results showed only modest global methylation differences (less than 5%) between naïve and primed states, but instead suggested significant demethylation typically tend to distribute near specific regions where other regulatory elements are also clustered. Moreover, notable methylation landscape changes are observed during retinal differentiation in both global and local scales, increasing the complexity of epigenetic regulation during retinal lineage specification and leading to further need of multilayer molecular analysis. By revealing these methylation and other related regulation dynamics, this study aims to provide new understanding and implications of epigenetic regulatory networks in stem cell pluripotency control and differentiation efficiency.

# Methods

## Cell culture

### Undifferentiated cell culture

All cell lines used in this study were provided and cultured by Dr. Zambidis lab, including two hESC lines (H9 and RUES02) and one hiPSC line (E5C3). They were cultured under two conditions: conventional Essential 8 (E8) medium and a chemically reverted naïve state using the LIF5i to LIF-3i tankyrase-inhibited (TIRN) protocol, as previously outlined17–19​. Briefly, primed hPSC cultures were expanded in a lab-prepared version of E8 medium, consisting of DMEM/F-12 with 2.5 mM L-Glutamine, 15 mM HEPES, and 14 mM sodium bicarbonate (ThermoFisher Scientific, Cat# 11330). This was further supplemented with 50–100 ng/mL recombinant human FGF-basic (Peprotech), 2 ng/mL recombinant human TGF-β1 (Peprotech), 64 µg/mL L-ascorbic acid 2-phosphate magnesium (Sigma), 14 ng/mL sodium selenite (Sigma), 10.7 µg/mL recombinant human transferrin (Sigma), and 20 µg/mL recombinant human insulin (Peprotech). These primed cultures were maintained on Vitronectin XF-coated 6-well tissue culture plates (STEMCELL Technologies, Corning) and passaged every 5–6 days using non-enzymatic dissociation reagents such as Versene solution or enzyme-free PBS-based dissociation buffer (ThermoFisher Scientific, #13151).

For naïve conversion, hPSCs were chemically reverted using the TIRN method described by Zimmerlin et al. After an initial adaptation step in LIF-5i medium, the cultures were propagated in LIF-3i medium17,19. This medium is based on DMEM/F-12 supplemented with 20% KnockOut Serum Replacement (ThermoFisher Scientific), 0.1 mM MEM Non-Essential Amino Acids (ThermoFisher Scientific), 1 mM L-Glutamine (ThermoFisher Scientific), 0.1 mM β-mercaptoethanol (Sigma), 20 ng/mL recombinant human LIF (Peprotech), 3 µM CHIR99021 (Peprotech), 1 µM PD0325901 (Sigma or Peprotech), and 4 µM XAV939 (Peprotech). Reversion was initiated by transferring 4–5-day-old primed hPSC cultures from E8 into LIF-5i medium supplemented with 10 µM Forskolin, 2 µM purmorphamine, and 10 ng/mL recombinant human FGF2 (all from Peprotech) based on LIF-3i medium. After 12–24 hours, cells were enzymatically dissociated using Accutase (ThermoFisher Scientific) and plated onto irradiated mouse embryonic fibroblast (MEF) feeders in LIF-5i medium for 3–5 days. Following this, all maintenance and passaging were done in LIF-3i medium on MEFs, using single-cell dissociation with Accutase.

### Retinal organoid culture

Only hESC lines (H9 and RUES02) were used for retinal organoid development. Naïve hPSCs were dissociated using Accutase (STEMCELL Technologies) and subjected to a 30–45 minute pre-plating step on gelatin-coated plates to eliminate residual mouse embryonic fibroblast (MEF) feeders. In parallel, primed hPSCs were detached using Versene (ThermoFisher Scientific). Both naïve and primed hPSCs were counted using the Countess automated cell counter (ThermoFisher Scientific) after Trypan Blue staining (ThermoFisher Scientific). Retinal differentiation was conducted based on previously established protocols with slight modifications20–23. Briefly, embryoid bodies (EBs) were generated using 1.2 × 10⁶ cells per well in AggreWell 400 plates (STEMCELL Technologies), following the manufacturer’s instructions. After 24 hours, EBs were harvested and transferred to low-attachment 6-well plates for suspension culture over a 6-day period. During this time, the culture medium was gradually transitioned from either E8 or L3i medium to Neural Medium 1, composed of DMEM/F12 (1:1), 1% N2 Supplement (ThermoFisher Scientific), 1× MEM Non-Essential Amino Acids (NEAA; ThermoFisher Scientific), and 2 µg/mL heparin (STEMCELL Technologies). On day 7, EBs were plated onto growth factor-reduced Matrigel-coated plates (Corning) at a density of 20 EBs per cm² in Neural Medium 1. By day 16, the medium was switched to Neural Medium 2, consisting of DMEM/F12 (3:1), 2% B27 supplement without vitamin A (ThermoFisher Scientific), and 1× NEAA. Until day 16, the medium was refreshed every other day, after which daily changes were performed.

During the third and fourth weeks of differentiation, horseshoe-shaped structures and forebrain neurospheres were carefully dissected from adherent cultures using a 0.125 mm tungsten needle (Roboz Surgical) under a sterile stem cell workstation. These were transferred to low-attachment 6-well plates for suspension culture in Neural Medium 3, composed of DMEM/F12 (3:1), 2% B27, 1× NEAA, and 1% antibiotic-antimycotic. Within 1–2 days of dissection, the horseshoe structures developed into three-dimensional retinal cups (RCs), while the forebrain neurospheres formed optic vesicles. The culture medium was then changed three times per week. Depending on the timing of horseshoe and neurosphere dissection (typically between weeks 3 and 5), cultures were transitioned to Neural Medium 4 approximately 1–1.5 weeks post-dissection. This medium included DMEM/F12 (3:1) with GlutaMAX (ThermoFisher Scientific), 2% B27, 1× NEAA, 10% fetal bovine serum, 100 µM taurine (Sigma), and 1% antibiotic-antimycotic. Retinal cups were further treated with 1 µM all-trans retinoic acid (Sigma) from weeks 10 to 14, followed by 0.5 µM for the remainder of the culture period.

## Bioinformatics

### Bulk RNA-Seq studies

Samples were processed as described above, washed with PBS, and snap-frozen in pellets containing 1–2 million cells. mRNA isolation and strand-specific library preparation were performed by the Genetic Resources Core Facility at the Johns Hopkins Department of Genetic Medicine. Specifically, polyadenylated mRNA was enriched using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB#E7490), and libraries were generated using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB#E7760). The resulting stranded libraries were sequenced on an Illumina NovaSeq 6000 platform using 50 bp paired-end, dual-indexed reads.

The raw RNA-seq read alignment was carried out using STAR (v2.7.10a)24, with genome indexing based on the Homo\_sapiens.GRCh38.dna.primary\_assembly.fa FASTA file and Homo\_sapiens.GRCh38.111.gtf annotation file from Ensembl v111. STAR was run with the following parameters: --outSAMtype BAM Unsorted SortedByCoordinate and --quantMode TranscriptomeSAM GeneCounts. The BAM files were transformed into BigWig files for IGV visualization using deepTools (version 3.5.4) with the following commands: bamCoverage -b sorted.bam -o output.bw --normalizeUsing CPM25. To quantify gene expression, summarizeOverlaps from the R packages ‘Rsamtools’, ‘GenomicFeatures’, and ‘GenomicAlignments’ was used, with settings: features = exonsByGene, reads = bamfiles, mode = "Union", singleEnd = FALSE, ignore.strand = FALSE, and fragments = TRUE. Differential expression analysis was performed using DESeq226. Ensembl gene IDs were converted to gene symbols using the AnnotationHub package and the current EnsDb.Hsapiens annotation objects (e.g., v111). Volcano plots were created using the ‘EnhancedVolcano R’ package. For heatmap construction, we used a matrix of mean-subtracted regularized log-transformed counts from DESeq2. Batch effects were corrected using the removeBatchEffect function from ‘limma’27, and final heatmaps were visualized with the ‘ComplexHeatmap’ package28. Transcription factor annotations for volcano plots and heatmaps were derived from published data of the undifferentiated state and RO29,30. For gene set enrichment analysis (GSEA), R packages ‘clusterProfiler’ and ‘ReactomePA’ were used with function gsePathway(geneList\_KEGG, eps = 0, pvalueCutoff = 0.05, pAdjustMethod = "BH", verbose = FALSE)31,32.

### ChIP-Seq studies

ChIP-seq was conducted using only primed and TIRN-reverted RUES02 hESC lines in the undifferentiated state. For each condition, 100 million cells were crosslinked using a freshly prepared fixation solution containing 11% formaldehyde, 5 M NaCl, 0.5 M EDTA (pH 8.0), and 1 M HEPES (pH 7.9), for 15 minutes at room temperature with gentle agitation. Fixation was quenched by adding 1/20 volume of 2.5 M glycine. From this point forward, samples were kept on ice or refrigerated, washed twice with cold PBS containing 0.5% Igepal, and supplemented with 1 mM PMSF in the second wash before snap-freezing the resulting pellets. Subsequent steps—including chromatin shearing, immunoprecipitation, qPCR validation, library construction, quality control, barcoding, and sequencing—were performed by Active Motif. Duplicate ChIP samples from both primed and TIRN hESC conditions were processed using the following antibodies: H3K4me3, H3K27ac, H3K27me3, PARP1, NANOG, POU5F1 (OCT4), and SOX2 (Active Motif). Input control libraries were created by pooling equal amounts of genomic DNA from all unprecipitated samples. Sequencing was performed on an Illumina NextSeq 500 using 75-nt single-end reads, and alignments were conducted using the VWA algorithm against the hg38 human genome reference.

Post-alignment processing involved sorting BAM files using the ‘sort’ function in samtools33, filtering uniquely aligned reads using view -b -F4 and view -b -q25, and removing PCR duplicates with ‘markdup’. Peak calling was initially performed using MACS234. To estimate statistical cutoffs, macs2 callpeak was run with the --cutoff-analysis option and the following parameters: -f BAM -B --gsize=hs --tsize=75 --bw=200 -m 5 50 -n test\_p. Final peak calling was executed with: callpeak -f BAM -B --gsize=hs --tsize=75 --bw=200 -m 5. For reproducibility assessment using the Irreproducible Discovery Rate (IDR) method34, low-stringency p-value thresholds were used for specific targets: H3K4me3 (p = 0.001), NANOG (p = 0.01), PARP1 (p = 0.05), POU5F1 (p = 0.01), and SOX2 (p = 0.001). The resulting narrowPeak files were ranked by -log10(p-value) using sort -k8,8nr, and replicates were combined with the idr --rank p.value command. For broader marks such as H3K27me3, and H3K27ac, a p-value cutoff of 0.0001 was applied, with MACS2 parameters including: callpeak -f BAM -B --SPMR --broad --broad-cutoff 0.001 --gsize=hs --tsize=75 --bw=200 -m 5 50. Consensus broad peaks were identified using the bedtools intersect command 35.

ChIP-seq data for human DUX4 (SRA accession: PRJNA377315) was retrieved using the SRA Toolkit prefetch utility36. Paired-end FASTQ files were aligned to hg38 using bowtie2 with the original study’s parameters: -t --sensitive-local -p 20 --no-mixed --no-discordant37. SAM files were converted to BAM and sorted with samtools ‘view’ and ‘sort’33. Reads were filtered for mapping quality (-q10), and pairing statistics were confirmed using ‘flagstat’. Peak calling was performed using MACS2 with parameters: callpeak -f BAMPE -B --SPMR -q 0.05 --gsize=hs38. IDR was again used to define consensus peaks34. For all ChIP-seq profiles, BAM files were indexed using samtools ‘index’, and normalized signal tracks were generated using deepTools bamCoverage with RPGC normalization: bamCoverage --binSize 10 --normalizeUsing RPGC --smoothLength 60 --extendReads 75 --centerReads --numberOfProcessors 20 --effectiveGenomeSize 2913022398. To generate input-subtracted tracks for IGV visualization, we used the deepTools ‘bigwigCompare’ function, and average bigWig tracks across replicates were calculated using ‘bigwigAverage’25.

To identify differentially-bound regions for NANOG, OCT4, and SOX2, we used the ‘DiffBind’ Bioconductor R package. Replicate IDR peaksets were used to generate DBA objects (DiffBind::dba) for each transcription factor and for the combined set. Read counts from BAM files were quantified with dba.count, with peaks recentered to a uniform width of 400 bp. Data normalization based on sequencing depth was performed with dba.normalize, and primed hESC samples were set as the reference group using dba.contrast(minMembers = 2). Differential analysis was performed with DESeq2 (dba.analyze)26. Peaks overlapping blacklisted regions in the hg38 genome were excluded, and ChIP input control files were used to define greylists. All bound intervals were retrieved using dba.report(method = DBA\_DESEQ2, contrast = 1 or 4, bCounts = TRUE, bCalled = TRUE, th = 1) depending on whether individual or SON co-binding analysis was performed. Genomic annotations were added using the ‘ChIPseeker’ package with the ‘annotatePeak’ function and an Ensembl v111 EnsDb object from AnnotationHub39.

### WGBS studies

WGBS library preparation and sequencing were performed using previously established protocols in the Genetic Resources Core Facility at the Johns Hopkins Department of Genetic Medicine. Briefly, libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs). For each sample, 500 ng of genomic DNA (gDNA) was quantified using the Qubit dsDNA BR Assay Kit (Invitrogen) and spiked with 1% unmethylated Lambda DNA (Promega, Cat# D1521) to assess bisulfite conversion efficiency. DNA was sheared to an average fragment size of approximately 350 bp using Covaris S220 and LE220 Focused-ultrasonicator instruments. Size selection was performed using AMPure XP magnetic beads, and fragments between 300–400 bp were retained for downstream processing. Following size selection, bisulfite conversion was conducted using either the EZ DNA Methylation-Gold Kit or the EZ DNA Methylation-Lightning Kit (Zymo Research, Cat# D5030), according to the manufacturer’s instructions. After bisulfite treatment, DNA libraries were amplified using the KAPA HiFi Uracil+ polymerase (Kapa Biosystems, Cat# KK282) with the following thermal cycling program: initial denaturation at 98°C for 45 seconds, followed by 8 cycles of 98°C for 15 seconds, 65°C for 30 seconds, and 72°C for 30 seconds, with a final extension at 72°C for 1 minute. The resulting libraries were cleaned with AMPure XP beads and assessed using the Agilent 2100 Bioanalyzer with the High-Sensitivity DNA assay. Bioanalyzer QC was also conducted on samples after shearing and size selection. Library quantification was performed using qPCR with the KAPA Library Quantification Kit for Illumina platforms (Kapa Biosystems, Cat# KK4824) on a CFX384 Real-Time PCR Detection System (Bio-Rad). Final WGBS libraries were sequenced on an Illumina HiSeq 4000 platform using 150 bp paired-end reads. A 5% PhiX control library (Illumina), not indexed, was spiked in as a sequencing control.

WGBS reads were trimmed using ‘trim\_galore’ (version 0.6.10) with the following parameters: trim\_galore --paired –fastqc forward.fastq.gz backward.fastq.gz40​. The alignment was performed using ‘Bismark’ (version 0.24.2)41. The Bismark library was prepared with the human hg38 Ensembl v111 Homo\_sapiens.GRCh38.dna.primary\_assembly.fa file using the parameters: bismark\_genome\_preparation --bowtie2 folder\_containing\_reference\_file. Read alignment was performed using ‘Bismark’ and ‘Bowtie2’ (version 2.5.3) using the options: bismark --bam --bowtie2 --genome reference\_folder​ -1 trimmed\_forward.fastq.gz -2 trimmed\_backward.fastq.gz. Multiple bam files from the same sample were merged during deduplication using Bismark with the options: deduplicate\_bismark -p --bam --multiple <lane1.bam> <lane2.bam>. The files for downstream analysis were produced by Bismark methylation extraction with the parameters: bismark\_methylation\_extractor -p --cytosine\_report --no\_overlap –comprehensive --bedGraph --ucsc --genome\_folder reference\_folder deduplicated.bam​. The bedGraph files were transformed into BigWig formats using ‘bedGraphToBigWig’ (version 2.10) with the commands: bedGraphToBigWig in.bedGraph chrom.sizes out.bw (chrom.sizes files were also produced by Bismark). The BigWig files were used along with the SOX2, OCT4, NANOG differentially bound site data (in Bed-like format) to produce profile plots using deepTools (version 3.5.4) with the following code: computeMatrix reference-point --referencePoint center -b 1500 -a 1500 -R gain.bed loss.bed -S TIRN1.bw TIRN2.bw TIRN3.bw Primed1.bw Primed2.bw Primed3.bw -o center.gz; computeMatrixOperations relabel -m center.gz -o center\_relabel.gz --groupLabels "gain" "loss" --sampleLabels "TIRN1" "TIRN2" "TIRN3" "Primed1" "Primed2" "Primed3"; plotProfile -m center\_relabel.gz -o center\_profile.png --colors red crimson salmon royalblue mediumblue deepskyblue --plotType lines --perGroup --plotHeight 10 --plotWidth 10 --legendLocation lower-left --plotTitle "Distribution of Methylation Percentage on ... Sites (All Undiff)" --yMin x –yMax y (colors for RO: --colors orange darkgoldenrod gold black dimgrey darkgrey)25.

The cytosine reports generated by Bismark were used for DMR analysis using R package ‘DMRichR’ (version 1.7.8)42. All analysis results, including raw DMR curves and DMR enrichment on CpG contexts and genic regions, were generated by ‘DMRichR’ with the master function: DM.R <- function(genome = "hg38", testCovariate = "Group", EnsDb = TRUE). To run the code, the working directory should contain a design matrix (sample\_info.xlsx) in addition to the unzipped cytosine reports (**Table 1**).

**Table 1.** Sample design matrix used in ‘DMRichR’ package

with the label of experimental group (L3i) required to be later than that of control group (E8) in alphabetical order.

|  |  |
| --- | --- |
| **Name** | **Group** |
| hESC-1 | E8 |
| hESC-3 | E8 |
| hiPSC-1 | E8 |
| N-hESC-1 | L3i |
| N-hESC-3 | L3i |
| N-hiPSC-1 | L3i |

The reference used in ‘DMRichR’ was updated by running trace(DMRichR::annotationDatabases, edit = TRUE) and changing the original code at line 151 from ‘genome == "hg38" ~ "AH83216" to genome == "hg38" ~ " AH116291"’ (AH116291 is the EnsDb v111 reference). The global methylation value of each sample was also provided by DM.R in Excel table and the visualization (barplot) was made separately. The DMR curves were modified to remove data points and reorder based on absolute difference value. After loading the data file from RData directory, running the following codes:

library(AnnotationHub)

library(DMRichR)

# Modify DMR

o <- order(abs(sigRegions$difference), decreasing = TRUE)

sigRe\_reorder <- sigRegions[o]

testCovariate = "Group"

ah <- AnnotationHub()

query(AnnotationHub(), c("EnsDb", "Homo sapiens"))

TxDb <- ah[["AH116291"]]

annoTrack <- GenomicRanges::GRangesList(CpGs = DMRichR::getCpGs("hg38"),

Exons = DMRichR::getExons(TxDb),

compress = FALSE)

pdf("Reorder.pdf", height = 4, width = 8)

tryCatch({

DMRichR::plotDMRs2(bs.filtered,

regions = sigRe\_reorder,

testCovariate = testCovariate,

extend = (end(sigRe\_reorder) - start(sigRe\_reorder) + 1)\*2,

addRegions = sigRe\_reorder,

annoTrack = annoTrack,

regionCol = "#FF00001A",

addPoints = FALSE,

lwd = 2,

qval = FALSE,

stat = FALSE,

horizLegend = FALSE)

},

error = function(error\_condition) {

print(glue::glue("Warning: One (or more) of your DMRs can't be plotted, \\

try again later by manually loading R Data and subsetting sigRegions"))

})

dev.off()

The statistics of all DMRs provided by DM.R were automatically annotated, and were imported into R to produce heatmap using ‘pheatmap’ and cross plot with bulk RNA-Seq data using ‘ggplot2’43. The genomic regions of the top DMRs with extension of two times of their sizes on upstream and downstream (same as DMR curve ranges) were used in IGV as the ranges for multiple track display44. The ChromHMM track was the “25 state-E008 H9 Cells” data downloaded from <https://egg2.wustl.edu/roadmap/web_portal/>45.

# Results

For clarification, TIRN and primed cells may be termed as L3i and E8 respectively in visualizations due to their culturing conditions. All RO data were sampled in week 12.

## TIRN-SC exhibited a modest global hypomethylation that was reduced or slightly reversed upon differentiation into retinal organoids.

To investigate the methylation pattern of functional human blastomere-like stem cells, we performed comparative whole genome bisulfite sequencing (WGBS) on three separate TIRN-reverted hESC and hiPSC lines and their corresponding non-reverted primed cell lines. Interestingly, the average global CpG methylation level of TIRN-SC lines did not show a significant difference (less than 5%) from that of their primed counterparts **(Fig. 1**), which was contrary to previous reports. Among the three cell lines, only one of them (RUES02) reached a ~5.5% methylation difference between TIRN and primed states, while the other two lines exhibited nearly the same methylation in the two states (**Fig. 2, Table 1**).

**Figure 1.** Average genome wide-CpG methylation levels of TIRN and primed cell lines in the undifferentiated state and RO

(n = 3 for each condition).

**Figure 2.** Genome wide-CpG methylation levels of individual cell lines

included in Fig.1.

**Table 2.** Genome wide-CpG methylation levels of individual cell lines

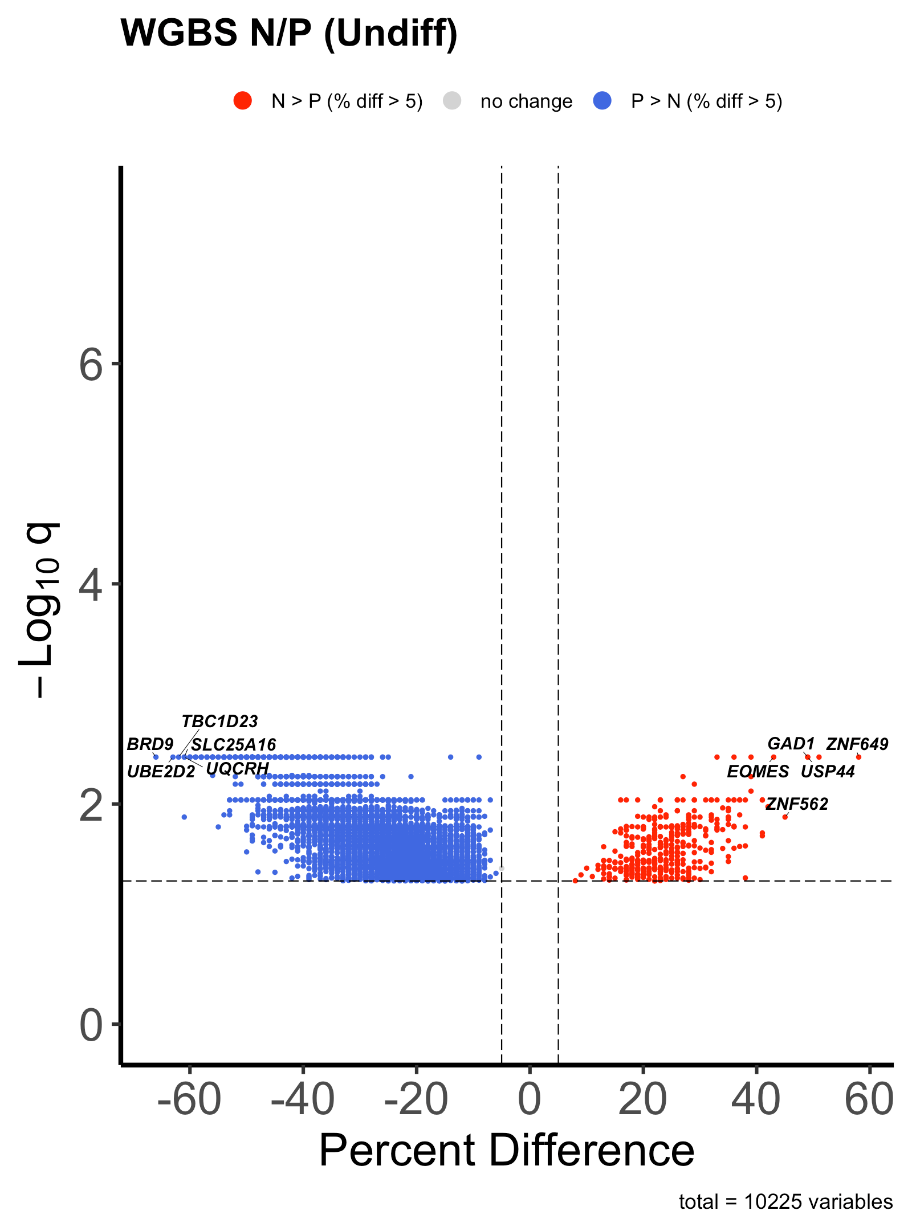
showed in Fig. 2.

|  |  |
| --- | --- |
| **Sample** | **Percent CpG methylation** |
| H9\_undiff\_TIRN​ | 75.70 |
| RUES02\_undiff\_TIRN​ | 75.74 |
| E5C3\_undiff\_TIRN​ | 78.36 |
| H9\_undiff\_primed​ | 77.17 |
| RUES02\_undiff\_primed​ | 81.19 |
| E5C3\_undiff\_primed​ | 77.93 |
| RUES02\_RO\_TIRN​ | 79.93 |
| H9\_RO\_TIRN1​ | 80.90 |
| H9\_RO\_TIRN2​ | 79.59 |
| RUES02\_RO\_primed​ | 79.83 |
| H9\_RO\_primed1​ | 73.99 |
| H9\_RO\_primed2​ | 76.62 |

We also performed WGBS on the retinal organoids (ROs) differentiated from TIRN-SC lines and their primed counterparts to track the methylation changes during differentiation. We observed a reversed result compared to the undifferentiated condition, with a modest hypermethylation (less than 5%) in the ROs from TIRN-SC compared to those from primed cell lines (**Fig. 1**). The methylation levels of the undifferentiated state and the ROs from both TIRN-SC and primed cell lines were all similar. Specifically, the RUES02 cell line with the largest difference in the undifferentiated condition completely lost that difference, and the modest hypomethylation in the H9 cell line in the undifferentiated condition was reversed to hypermethylation (**Fig. 2, Table 1**). Unfortunately, the RO data for the hiPSC line (E5C3) was not available for comparison.

## Methylation differences were significantly expanded at differentially methylated regions (DMRs) in the undifferentiated state.

Since the global methylation level did not show clear patterns between TIRN-SC and primed cell lines or between the undifferentiated state and the RO, we further performed differentially methylated region (DMR) analysis. A total of 10,225 DMRs of TIRN-SC versus primed cells were identified with a q-value less than 0.05 in the undifferentiated state, which were predominantly hypomethylated in TIRN-SC (**Fig. 3**), with the largest difference reaching more than 60%. These DMRs were widely distributed across the whole genome on various gene regions and with diverse CpG frequencies (**Fig. 4-6**). The methylation site enrichment analysis on CpG contexts revealed that the DMRs were more significantly enriched in CpG shores, slightly enriched in CpG islands and shelves, and depleted in open sea (**Fig. 7**), which was consistent with reports that most DMRs would occur in CpG shores46,47. Given that the majority of the DMRs were hypomethylated in TIRN-SC, the overall enrichment was almost the same as the hypomethylated DMR enrichment. Notably, although hypermethylated DMRs showed similar enrichment results in CpG shores, CpG shelves, and open sea, they were enriched in CpG islands at a significantly higher level, which may be associated with important methylation imprinting sites8,10. This enrichment analysis was also performed on genic regions. It indicated shared enrichment in promoters and 5’ UTRs and shared depletion in introns and intergenic regions in both hypermethylated and hypomethylated regions (**Fig. 8**), suggesting an overall tendency of clustering near regulatory regions. Compared to the similar global methylation level between TIRN-SC and primed cells, the significant methylation difference and clustering pattern of DMRs implied that the earlier differentiation state (blastomere-like) of TIRN-SC was likely induced by local demethylation primarily occurring at specific regulatory regions.

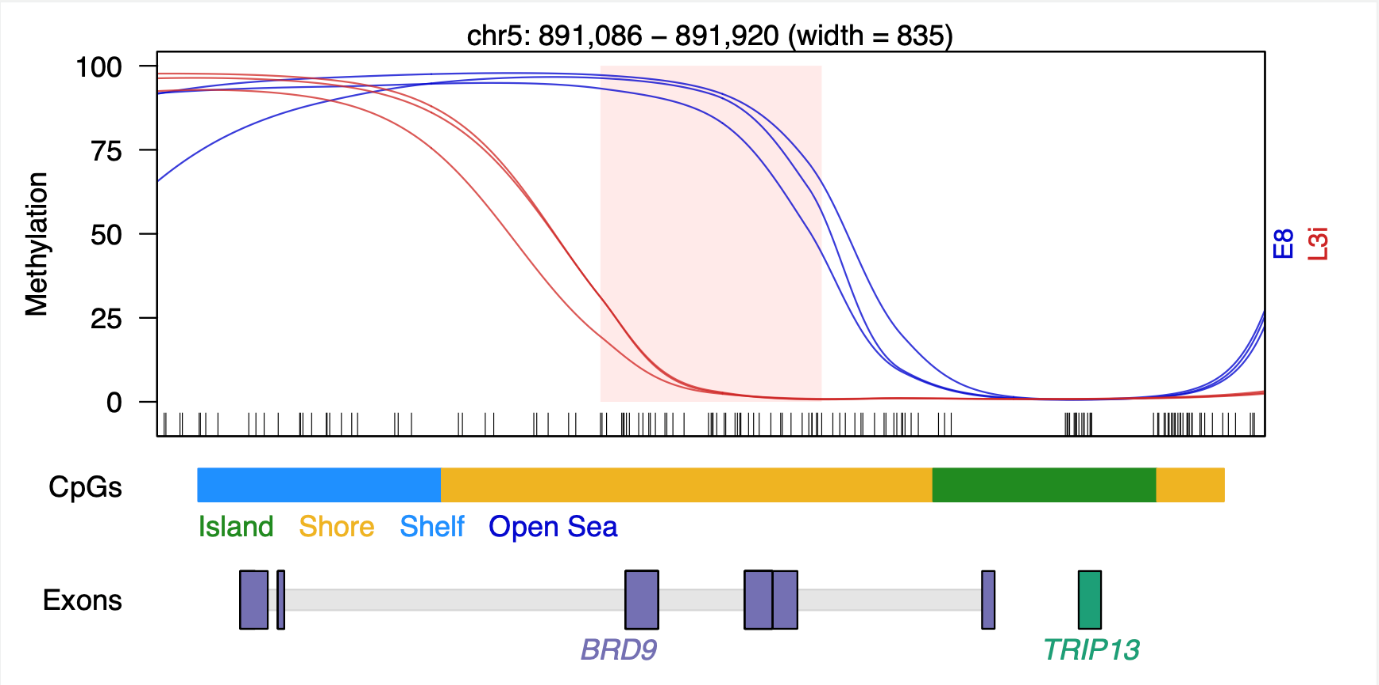


**9829**

**394**

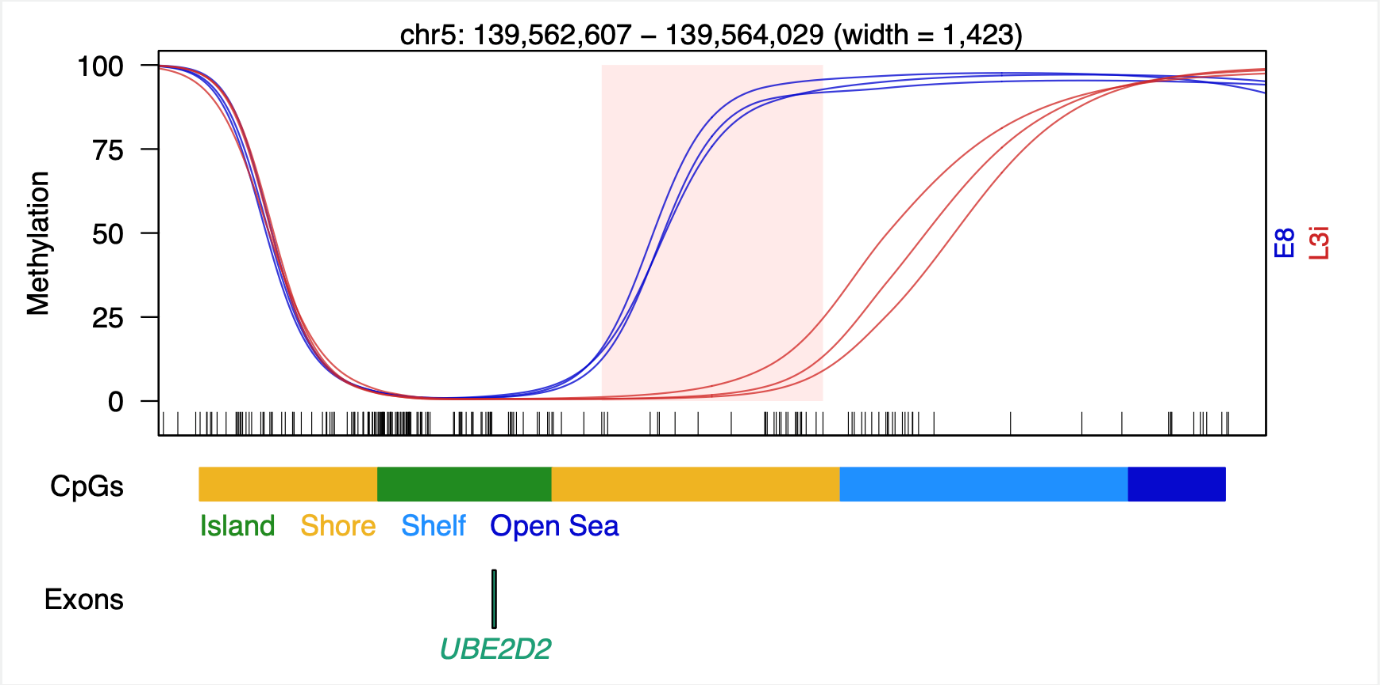
**Figure 3.** Volcano plot showing percent difference of CpG methylation of DMRs between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state

(WGBS, n = 3 independent cell lines). Colored dots show significant DMRs (|percent difference| > 5%, q < 0.05; delineated by dashed lines; number labeled on top) while non-significant ones are greyed out. DMRs are annotated with close genes, and the ones with top methylation differences are highlighted.



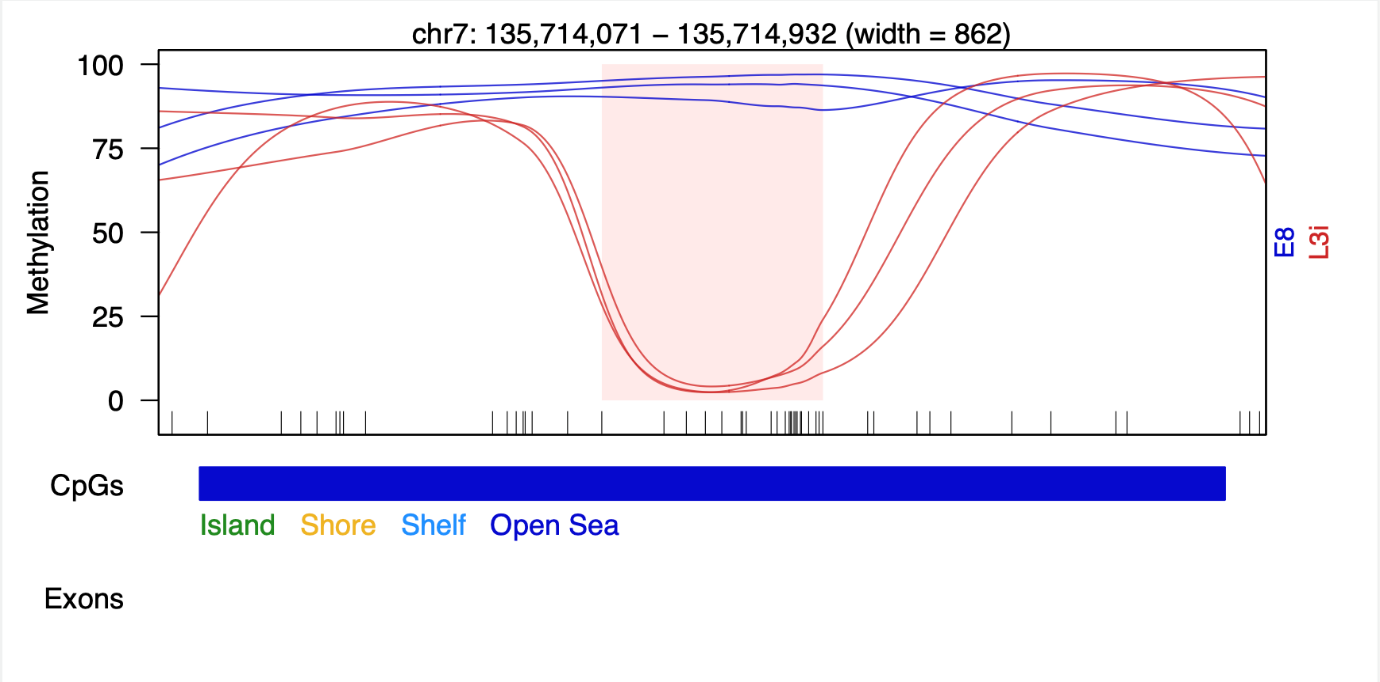
**Figure 4.** Smoothed methylation curve of sample DMR on exon and CpG shore between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state.

DMR is in the central region labeled by pink rectangle with 66% hypomethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).



**Figure 5.** Smoothed methylation curve of sample DMR on genic downstream and CpG shore between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state.

DMR is in the central region labeled by pink rectangle with 63% hypomethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).



**Figure 6.** Smoothed methylation curve of sample DMR on intergenic regions and open sea between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state.

DMR is in the central region labeled by pink rectangle with 60% hypomethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).

图表, 箱线图

AI 生成的内容可能不正确。

**Figure 7.** Enrichment analysis on CpG context of DMRs between TIRN-SC and primed hPSC in the undifferentiated state.

Fold enrichment is calculated as the odds ratio (> 1) or negative reciprocal of the odds ratio (< 1) of fisher exact test and the significance asterisk is labeled with q-value less than 0.05.

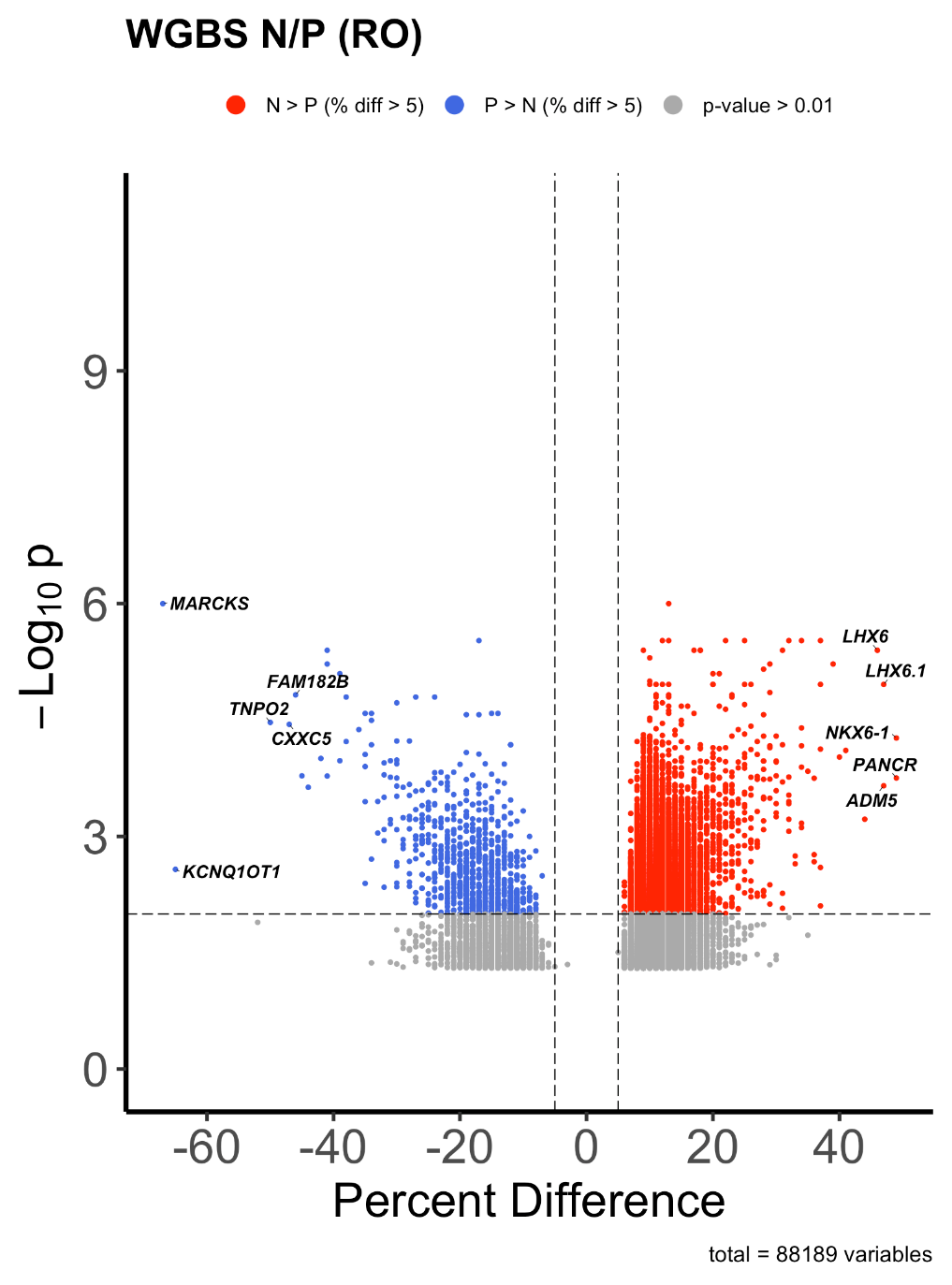
图表, 瀑布图

AI 生成的内容可能不正确。

**Figure 8.** Enrichment analysis on genic regions of DMRs between TIRN-SC and primed hPSC in the undifferentiated state.

Fold enrichment is calculated as the odds ratio (> 1) or negative reciprocal of the odds ratio (< 1) of fisher exact test and the significance asterisk is labeled with q-value less than 0.05.

The same DMR analysis on RO samples led to clearly different results. No DMR of TIRN-SC versus primed cells reached the q-value less than 0.05, therefore we reduced the threshold to a p-value less than 0.05 and identified 88,189 DMRs. Consistent with TIRN-SC modest global hypermethylation, the DMRs in RO were predominantly hypermethylated with lower overall divergence compared to the undifferentiated state despite the similar largest difference around 60% (**Fig. 9**). Also, the consistency of methylation level was reduced among samples, suggesting that some DMRs may be cell line-specific (**Fig. 10-12**). The methylation site enrichment on all CpG contexts in RO exhibited enrichment in CpG islands, shores, and shelves and depletion in open sea with overall lower fold enrichment compared to the undifferentiated state (**Fig. 13**). The overall higher enrichment on CpG islands compared to CpG shores reflected the condition in the predominant hypermethylated DMRs that showed this pattern in the undifferentiated state while the underrepresented hypomethylated DMRs kept their high enrichment in CpG shores. Such distinct enrichment patterns between hyper- and hypomethylated DMRs were also reflected in the analysis on genic regions where hypermethylated DMRs were modestly enriched in most regions except 3’ UTR and intergenic regions while the hypomethylated DMRs were depleted in most regions and only had significant enrichment in promoter regions (**Fig. 14**). These results suggested that the CpG methylation differences between TIRN-SC and primed cells were largely diminished in both global and local DMR scale after differentiation.



**883**

**22291**

**Figure 9.** Volcano plot showing percent difference of CpG methylation of DMRs between TIRN-SC (red) and primed (blue) hPSC in RO

(WGBS, n = 3 cell lines including one RUES02 and two H9). DMRs Colored dots show significant DMRs (|percent difference| > 5%, p < 0.01; delineated by dashed lines; number labeled on top) while non-significant ones are greyed out. DMRs are annotated with close genes, and the ones with top methylation differences are highlighted.

图片包含 图示

AI 生成的内容可能不正确。

**Figure 10**. Smoothed methylation curve of sample DMR on exon and CpG island between TIRN-SC (red) and primed (blue) hPSC in RO.

DMR is in the central region labeled by pink rectangle with 65% hypomethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).

图表

AI 生成的内容可能不正确。

**Figure 11.** Smoothed methylation curve of sample DMR on genic downstream and CpG island-shore combined between TIRN-SC (red) and primed (blue) hPSC in RO.

DMR is in the central region labeled by pink rectangle with 42% hypermethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).

图表

AI 生成的内容可能不正确。

**Figure 12.** Smoothed methylation curve of sample DMR on intergenic regions and open sea between TIRN-SC (red) and primed (blue) hPSC in RO.

DMR is in the central region labeled by pink rectangle with 38% hypermethylation, and the upstream and downstream extensions are double the size of the DMR. CpG contexts are labeled using color (island: green; orange: shore; light blue: shelf; dark blue: open sea).

图表, 瀑布图

AI 生成的内容可能不正确。

**Figure 13.** Enrichment analysis on CpG context of DMRs between TIRN-SC and primed hPSC in RO.

Fold enrichment is calculated as the odds ratio (> 1) or negative reciprocal of the odds ratio (< 1) of fisher exact test and the significance asterisk is labeled with q-value less than 0.05.

图表, 瀑布图, 箱线图

AI 生成的内容可能不正确。

**Figure 14.** Enrichment analysis on genic regions of DMRs between TIRN-SC and primed hPSC in RO.

Fold enrichment is calculated as the odds ratio (> 1) or negative reciprocal of the odds ratio (< 1) of fisher exact test and the significance asterisk is labeled with q-value less than 0.05.

## Local demethylation at core transcription factor binding sites drove the blastomere-like pluripotency of TIRN-SC.

Considering the pivotal roles of core transcription factors SOX2/OCT4/NANOG (SON) in embryonic stage pluripotency regulations12,13, we investigated the local CpG methylation level at the binding sites of these factors. Previous studies reported that these factors (SON) exhibit distinct binding patterns between naïve (including TIRN-SC) and primed pluripotent stem cells, reflecting their dynamic nature of pluripotency regulatory mechanisms48–50. Therefore, the genomic regions around chromatin immunoprecipitation sequencing (ChIP-seq) peaks with differential protein binding were specifically inspected for CpG methylation. These studies revealed that TIRN-SC exhibited significant local demethylation at the differential SON binding regions in the undifferentiated state (**Fig. 15**). Notably, only the differential binding regions gained in TIRN-SC showed such demethylation patterns, but the methylation levels at the regions lost in TIRN-SC were nearly the same for TIRN-SC and primed cells. The baseline analysis of methylation levels at random regions showed consistent results with the previous global methylation analysis (**Fig. 1, 1**6). When compared with the baseline conditions, methylation patterns at SON differential binding sites were clearly demonstrated, which started with baseline methylation levels at distal regions (~1.5 kb upstream and downstream) and gradually decreased to the lowest level when approaching the binding centers (**Fig. 15, 16**). The largest methylation differences between TIRN-SC and primed cells were also reached at the binding centers, with around 20% differences for all SON differential binding sites gained in TIRN-SC, strongly suggesting that CpG methylation may regulate the pluripotency of TIRN-SC via interactions with core transcription factors.

图表, 直方图

AI 生成的内容可能不正确。

**Figure 15.** Profile plot of WGBS methylation scores on the differential NANOG binding, SOX2 binding, OCT4 binding, and SON co-binding regions between TIRN-SC and primed hPSC in the undifferentiated state

(gain site: fold change > 2, FDR < 0.05; loss site: fold change < -2, FDR < 0.05).

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 16.** Profile plot of WGBS methylation scores on 100,000 random genomic regions (400 bp) between TIRN-SC and primed hPSC in the undifferentiated state

Due to the lack of RO ChIP-seq data, such analysis on RO was conducted by using the same differential binding sites from the undifferentiated state. The results again showed similar demethylation patterns around the SON binding centers, with reduced differences between TIRN-SC and primed cells, consistent with the global methylation and DMR analysis results (**Fig. 17, 18**). Specifically, the methylation differences at the binding sites of NANOG and OCT4 were nearly eliminated, while the SOX2 binding sites still maintained a low methylation difference, which could be supported by previous research revealing the downregulation of NANOG and SOX2 during retinal differentiation and the persistence of SOX2 due to its role in neural development51,52. The methylation at SON co-binding regions converged to similar levels after retinal differentiation, despite the significant differences in the undifferentiated state and the possible cell line variations (**Fig. 19**), showing the natural dynamics of methylation regulation at SON binding sites.

图表

AI 生成的内容可能不正确。

**Figure 17.** Profile plot of WGBS methylation scores on the differential NANOG binding, SOX2 binding, OCT4 binding, and SON co-binding regions between TIRN-SC and primed hPSC in RO

(gain site: fold change > 2, FDR < 0.05; loss site: fold change < -2, FDR < 0.05).

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 18.** Profile plot of WGBS methylation scores on 100,000 random genomic regions (400 bp) between TIRN-SC and primed hPSC in RO.

图表, 折线图, 直方图

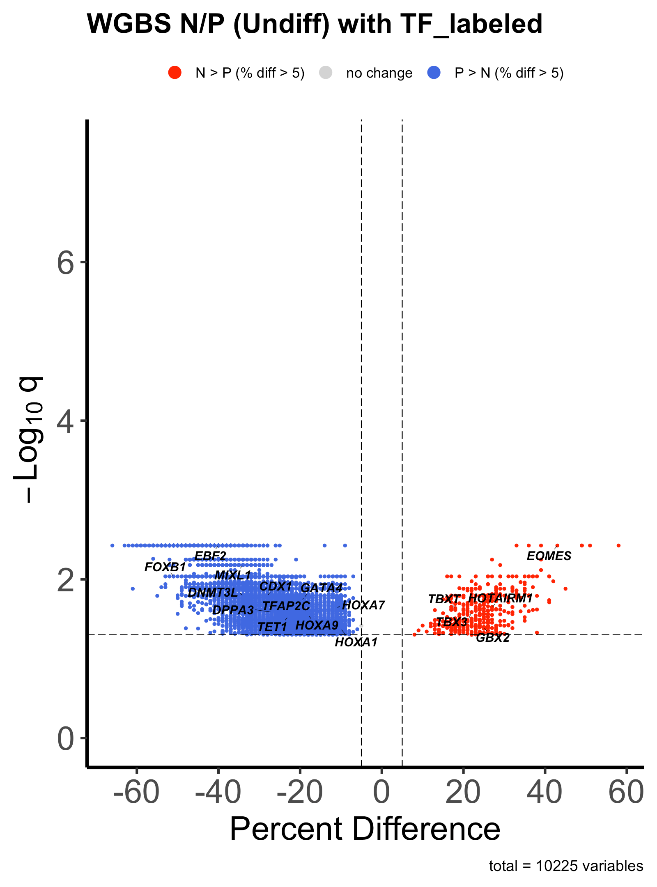
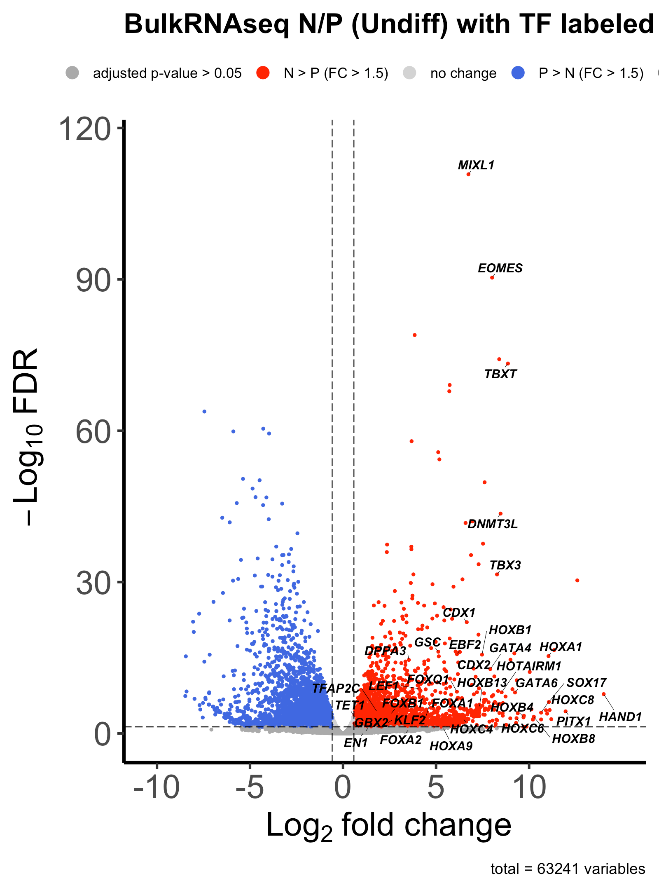
AI 生成的内容可能不正确。

**Figure 19.** Profile plot of WGBS methylation scores on the differential SON co-binding regions between TIRN-SC and primed hPSC during retinal differentiation

(gain site: fold change > 2, FDR < 0.05; loss site: fold change < -2, FDR < 0.05). Comparisons between the undifferentiated state and RO are only demonstrated for RUES02 and H9 cell lines due to lack of E5C3 RO data.

## The transcriptional programs of TIRN-SC were not directly regulated by CpG methylation.

It was recently reported that the TIRN-SC transcriptomics resembled human blastomere stage-specific transcriptional programs6. To verify such discovery and further explore its potential interplay with methylation, we combined such transcription and the new methylation data in the undifferentiated state to investigate the potential regulatory interactions between them. The blastomere-like phenotype of TIRN-SC was verified based on their simultaneous overexpression of transcription factors marking embryonic stages from pre-zygotic genome activation (pre-ZGA) to inner cell mass (ICM) compared to the primed cells (**Fig. 20, 21**), consistent with previous single-cell studies suggesting the heterogeneous transcriptome profiles of blastomere53,54. Gene set enrichment analysis (GSEA) also supported the earlier differentiation state of TIRN-SC compared to primed cells, with activated pathways related to maternal to zygotic transition, activation of HOX genes, and gastrulation (**Fig. 22**). Notably, the DNA methylation pathway was also activated, suggesting an interacting relationship between CpG methylation and transcription. However, this transcriptional signature was not directly regulated by CpG methylation. Although TIRN-SC exhibited a modest global hypomethylation, predominantly hypomethylated DMRs, and significant hypomethylation at SON binding sites, the overall numbers of up- and downregulated genes were quite close, and the methylation of those overexpressed transcription factors varied without clear patterns (**Fig. 20**). Also, the expected negative correlation between transcription and methylation could not be observed (**Fig. 23**).



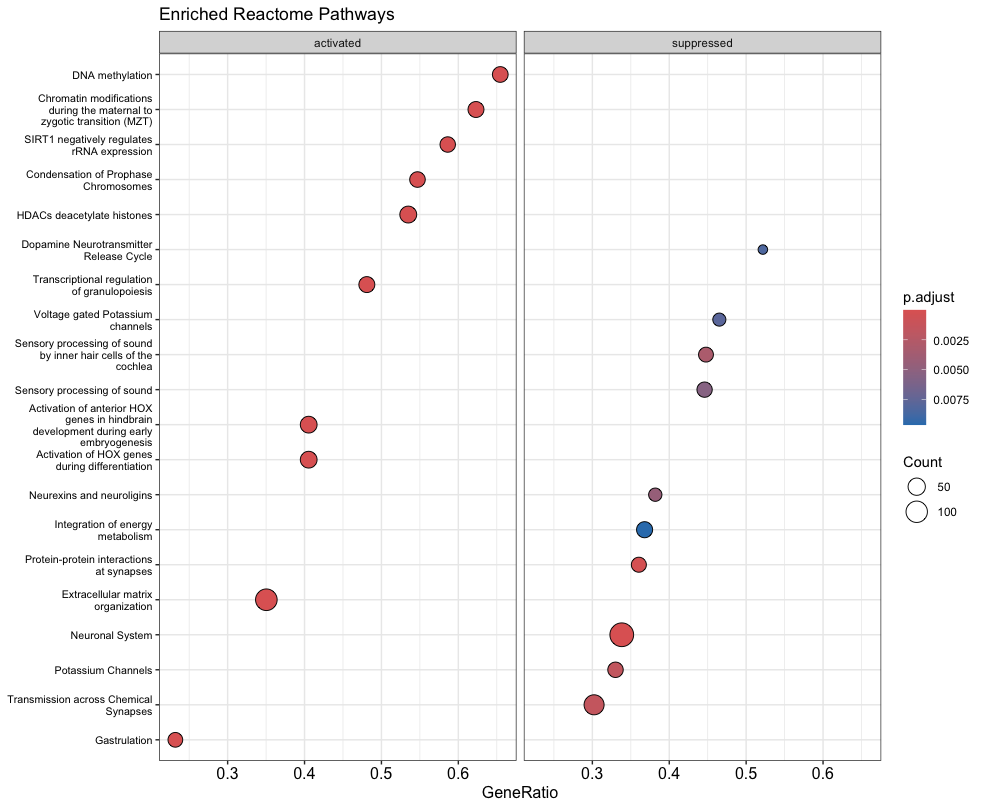
**Figure 20.** Volcano plot showing log2 fold change of RNA expression (**left**) percent difference of CpG methylation of DMRs (**right**) between TIRN-SC (red) and primed (blue) hPSC in the undifferentiated state

(WGBS, n = 3 independent cell lines). Colored dots show significant differentially expressed genes in left (|fold change| > 1.5, FDR < 0.05; delineated by dashed line); and significant DMRs in right (|percent difference| > 5%, q < 0.05; delineated by dashed lines) while non-significant ones are greyed out. Highlighted dots are key transcription factors expressed during human embryonic stages from pre-ZGA to ICM29.

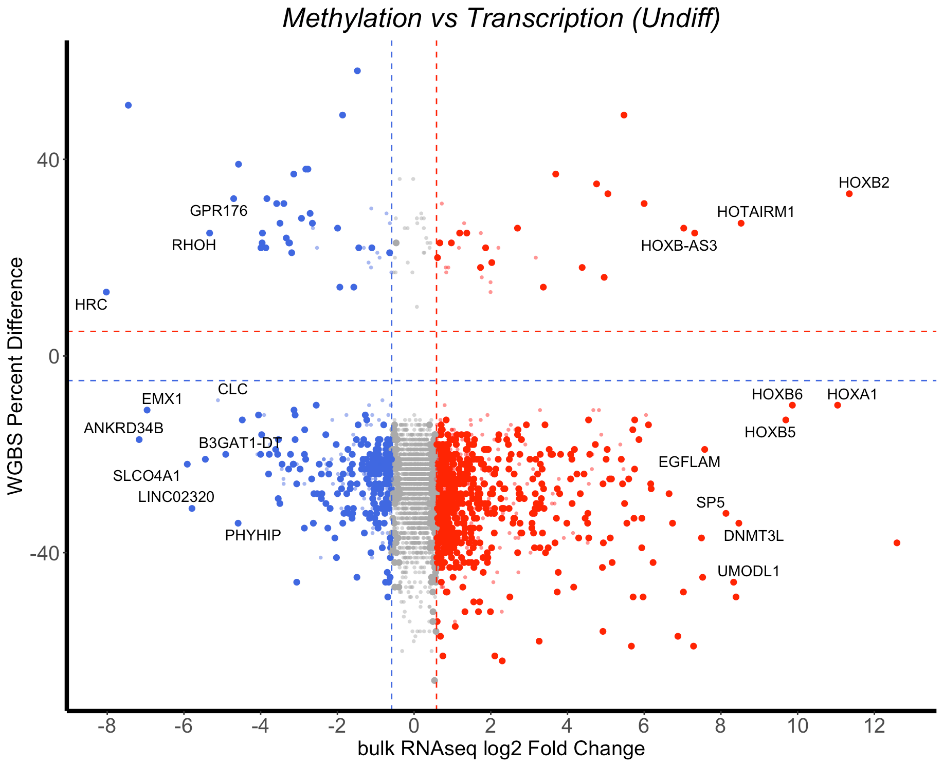


**Figure 21.** Heatmaps of RNA-seq data (FPKM) for TIRN-SC and primed hPSC lines comparing key transcription factors expressed during human embryonic stages from pre-ZGA to ICM

(FC>1.5, adjusted p-value <0.05)29.



**Figure 22.** Reactome GSEA pathway analysis of differential RNA expressions between TIRN and primed hPSC in the undifferentiated state.



**31**

**21**

**574**

**209**

**Figure 23.** Cross plot of log2 fold change of RNA-Seq (x-axis; RNA) vs percent difference of WGBS (y-axis; methylation) between TIRN and primed hPSC in the undifferentiated state.

Colored dots show significant differentially expressed genes (|fold change| > 1.5, FDR < 0.05; delineated by dashed line); and significant DMRs (|percent difference| > 5%, q < 0.05; delineated by dashed lines) with number labeled in corresponding quadrants. The non-significant ones are greyed out (q > 0.05) or with smaller size and higher transparency (q < 0.05; |fold change| > 1.5 for RNA-seq; |percent difference| < 5% and/or distance from TSS < 300bp for WGBS). Colors of dots are based on the overexpression of the genes in TIRN-SC (red) and primed hPSC (blue) and the ones with top RNA expression differences are highlighted.

The transcriptomics of TIRN-SC after retinal differentiation exhibited no evident bias towards overall up- or downregulation and no observable enrichment patterns for retinal transcription factors (**Fig. 24, 25**), which corresponded with their diminished methylation differences from the primed cells. Like conditions in the undifferentiated state, transcription factors were not clearly hyper- or hypomethylated in ROs differentiated from TIRN-SC, and no correlation between transcription and methylation was found (**Fig. 24, 26**). Surprisingly, despite the recent report suggesting better morphologies and improved maturation of TIRN-SC derived RO, GSEA analysis indicated no activated pathway related to retina but instead found phototransduction pathways suppressed along with the neural development pathways that were already suppressed in the undifferentiated state (**Fig. 22, 27**).

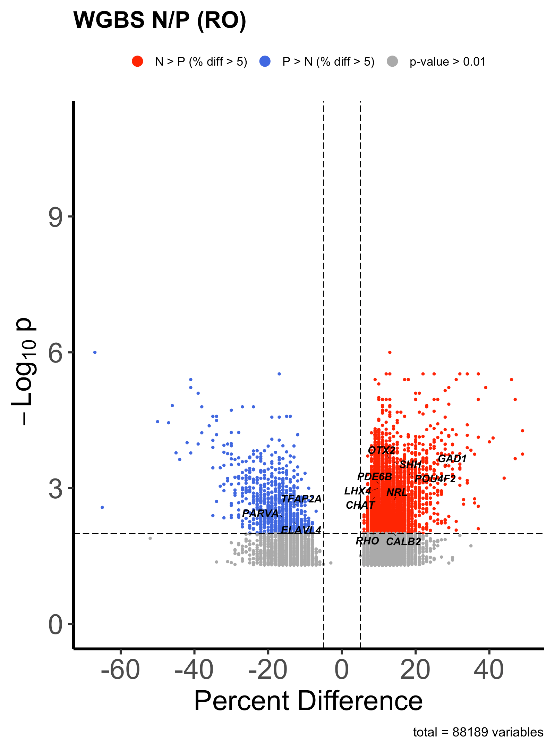
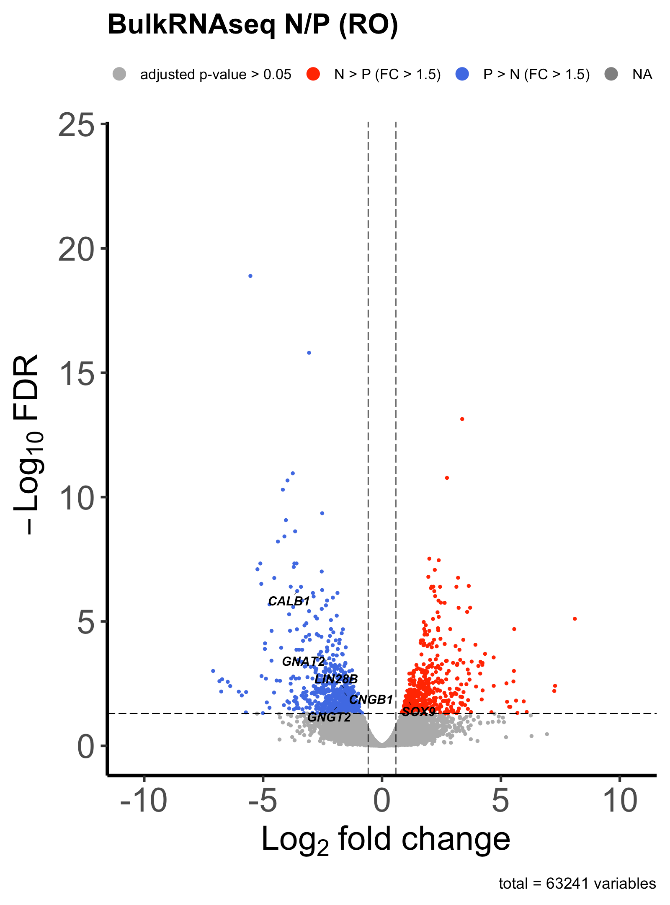
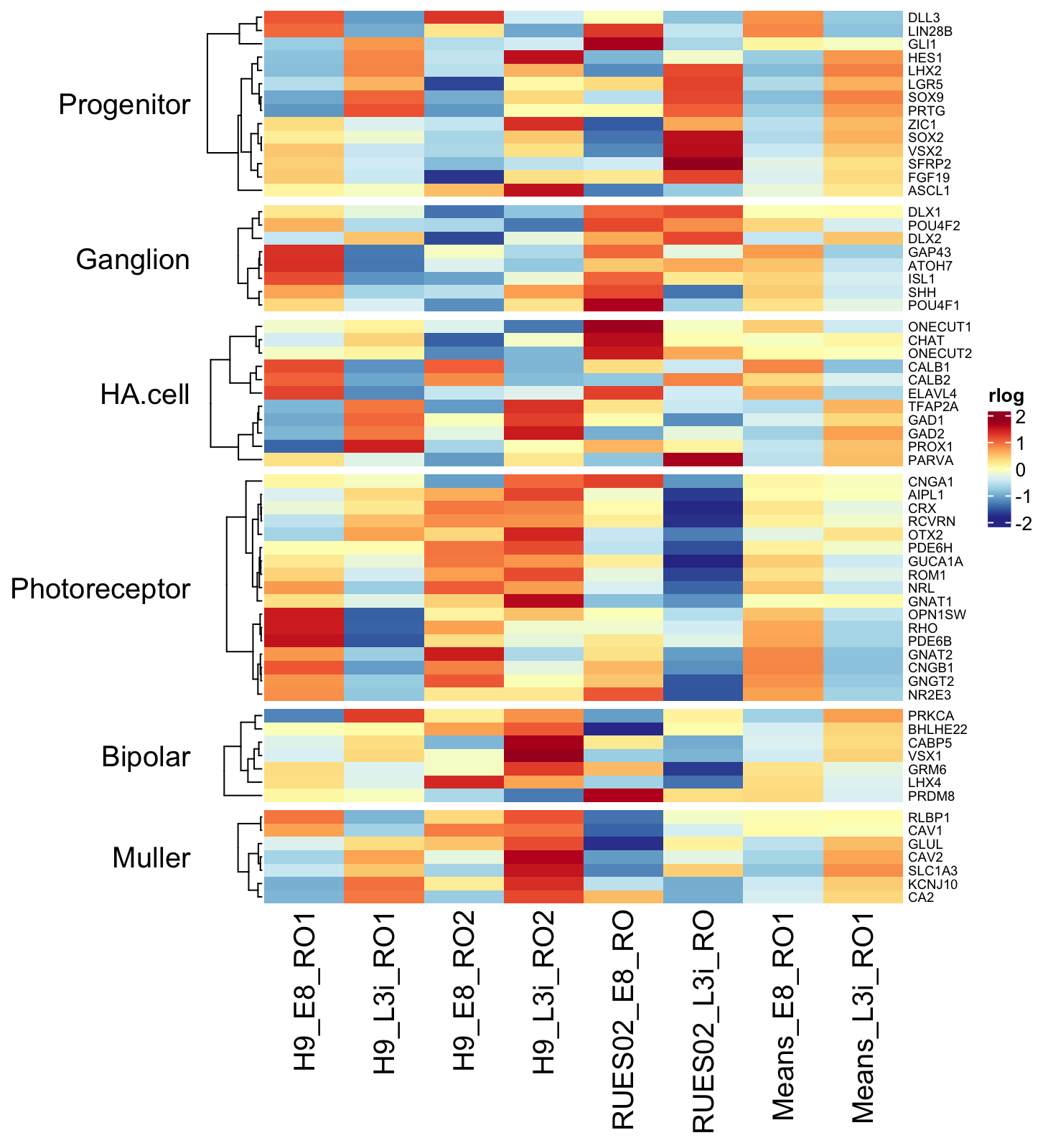


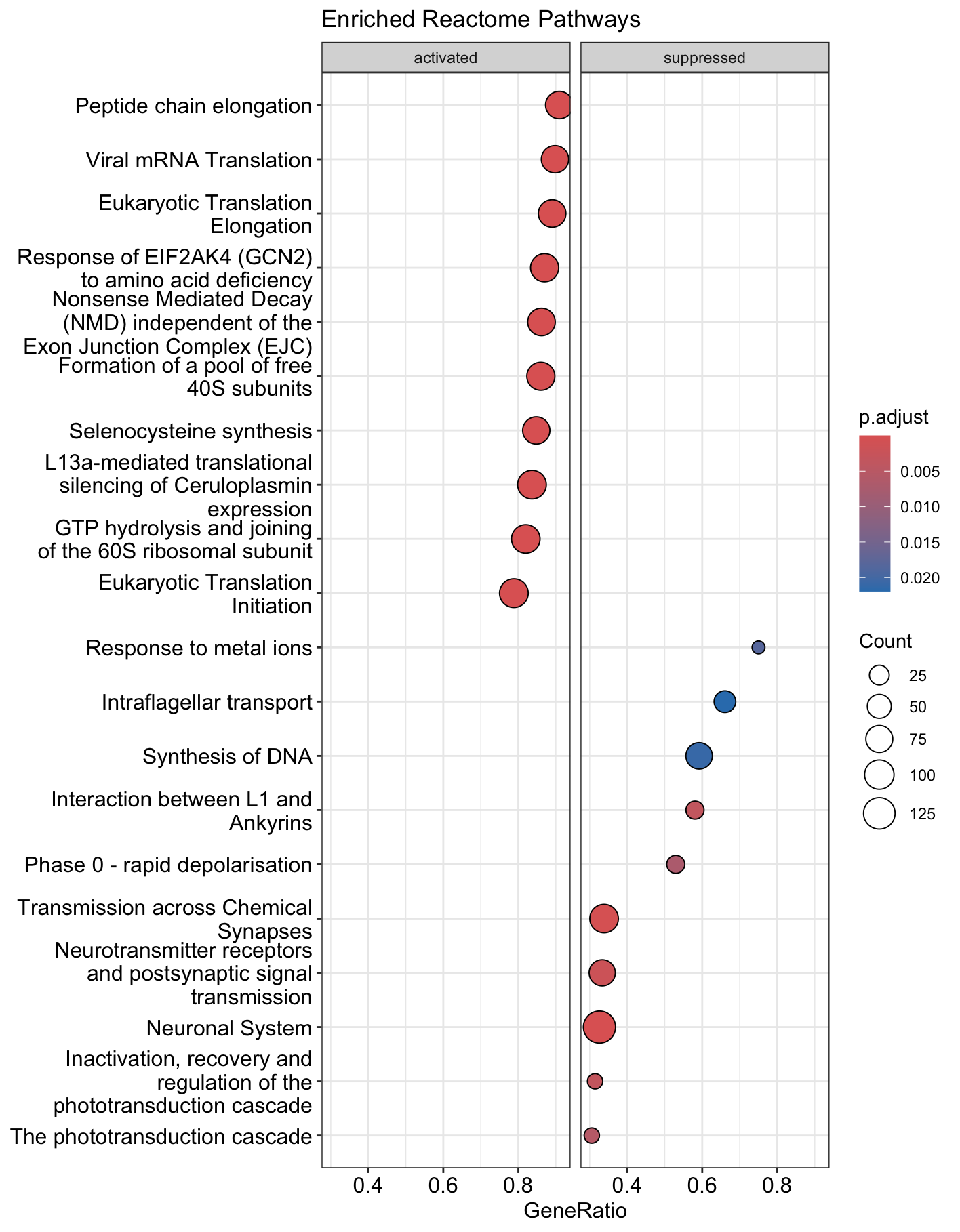
Figure 24. Volcano plot showing log2 fold change of RNA expression (**left**) percent difference of CpG methylation of DMRs (**right**) between TIRN-SC (red) and primed (blue) hPSC in RO

(WGBS, n = 3 independent cell lines). Colored dots show significant differentially expressed genes in left (|fold change| > 1.5, FDR < 0.05; delineated by dashed line); and significant DMRs in right (|percent difference| > 5%, q < 0.05; delineated by dashed lines) while non-significant ones are greyed out. Highlighted dots are key transcription factors expressed in human retinal differentiation30.

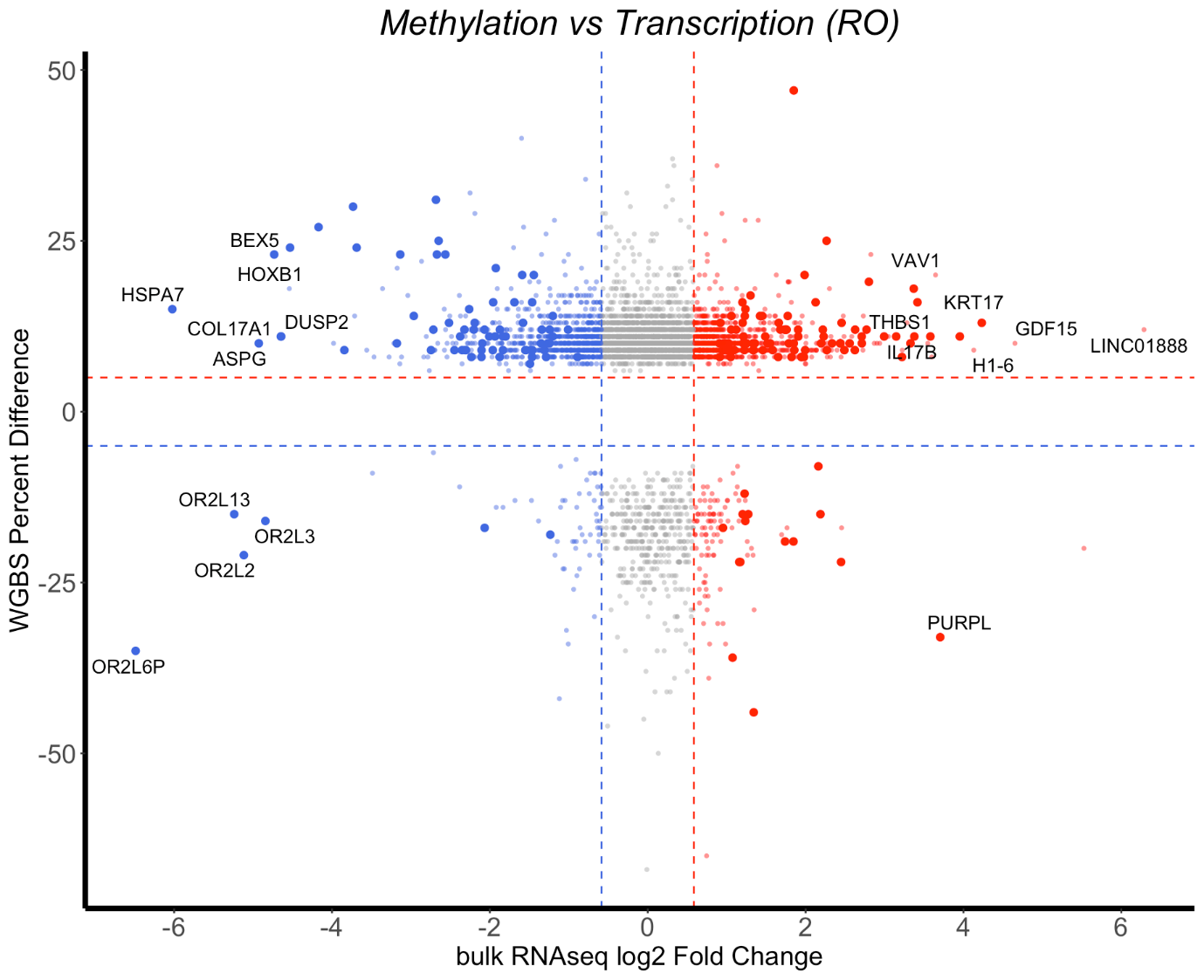


**Figure 25.** Heatmaps of RNA-seq data (FPKM) for TIRN-SC and primed hPSC lines comparing key transcription factors expressed in human retinal differentiation

(FC>1.5, adjusted p-value <0.05)30.



**Figure 26.** Reactome GSEA pathway analysis of differential RNA expressions between TIRN and primed hPSC in RO.

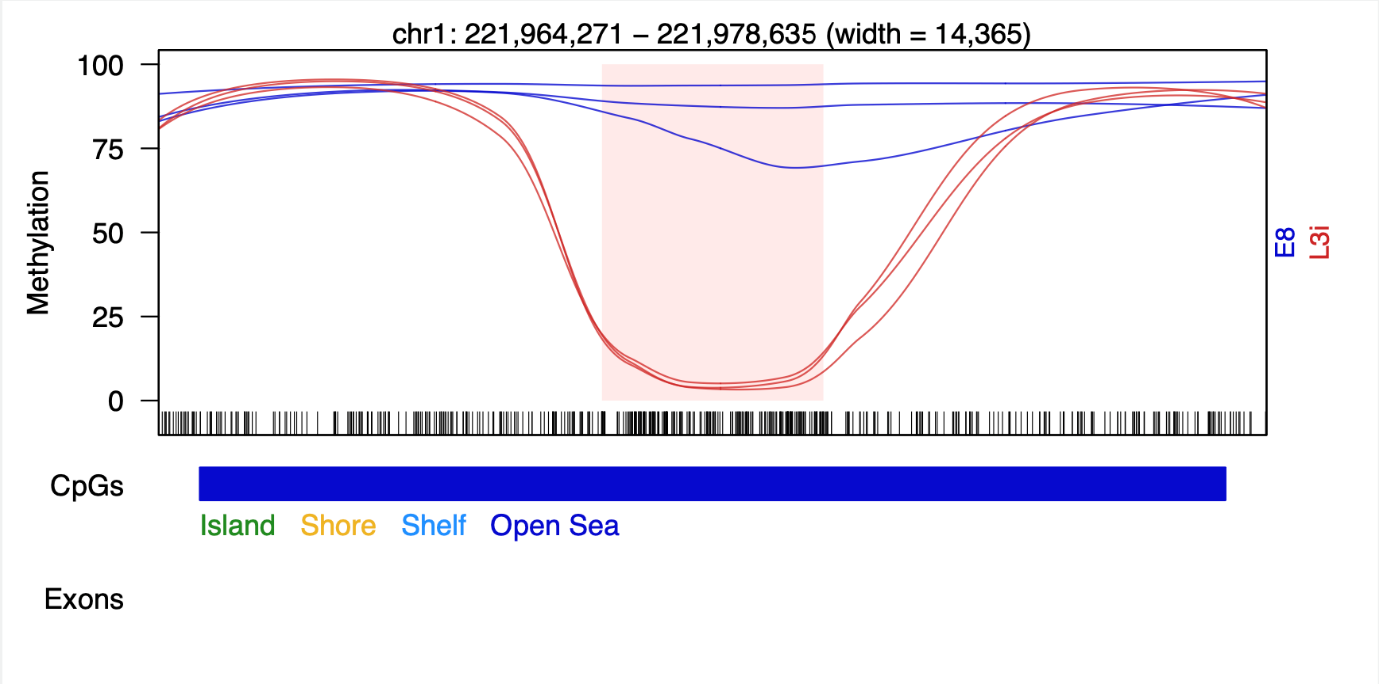


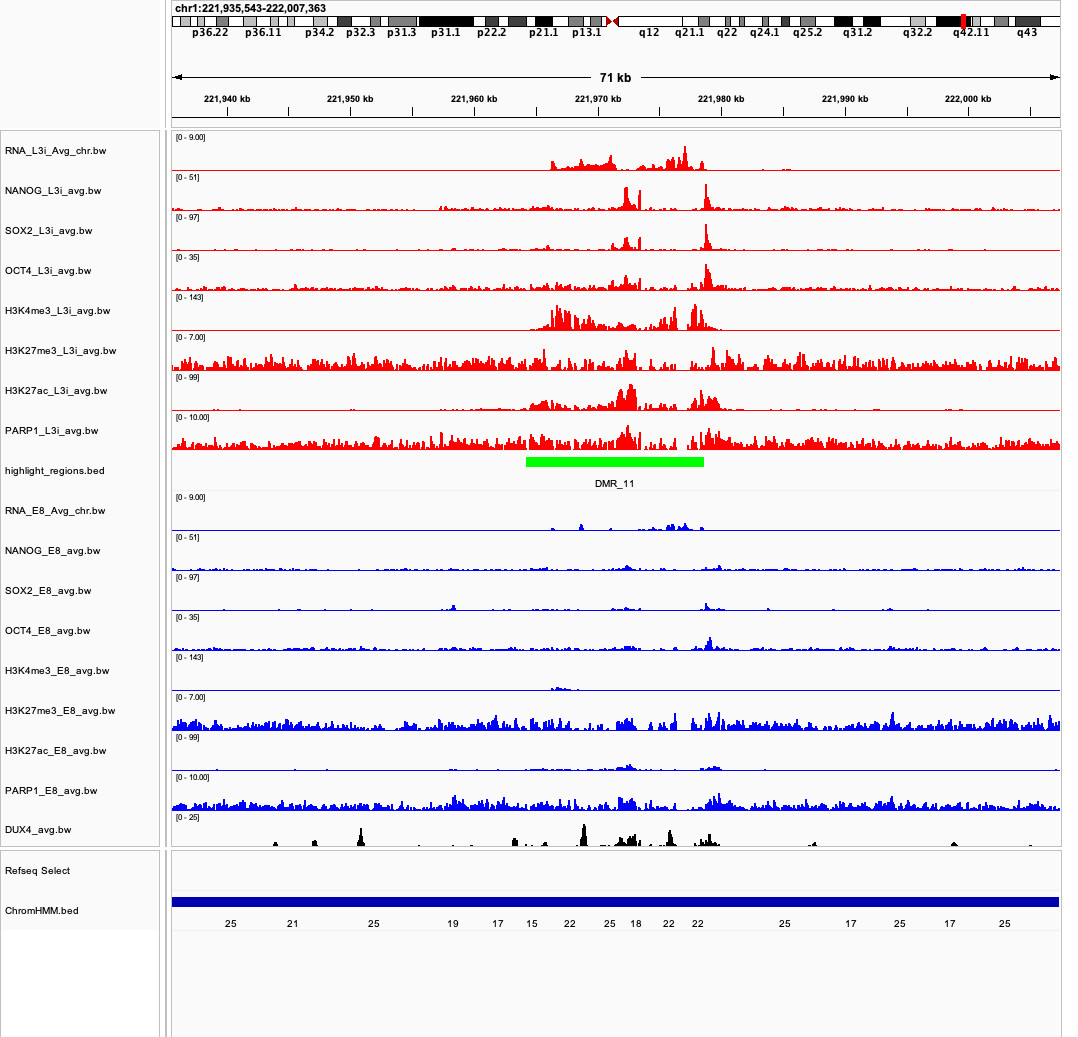
**Figure 27.** Cross plot of log2 fold change of RNA-Seq (x-axis; RNA) vs percent difference of WGBS (y-axis; methylation) between TIRN and primed hPSC in RO.

Colored dots show significant differentially expressed genes (|fold change| > 1.5, FDR < 0.05; delineated by dashed line); and significant DMRs (|percent difference| > 5%, q < 0.05; delineated by dashed lines) with number labeled in corresponding quadrants. The non-significant ones are greyed out (q > 0.05) or with smaller size and higher transparency (q < 0.05; |fold change| > 1.5 for RNA-seq; |percent difference| < 5% and/or distance from TSS < 300bp for WGBS). Colors of dots are based on the overexpression of the genes in TIRN-SC (red) and primed hPSC (blue) and the ones with top RNA expression differences are highlighted.

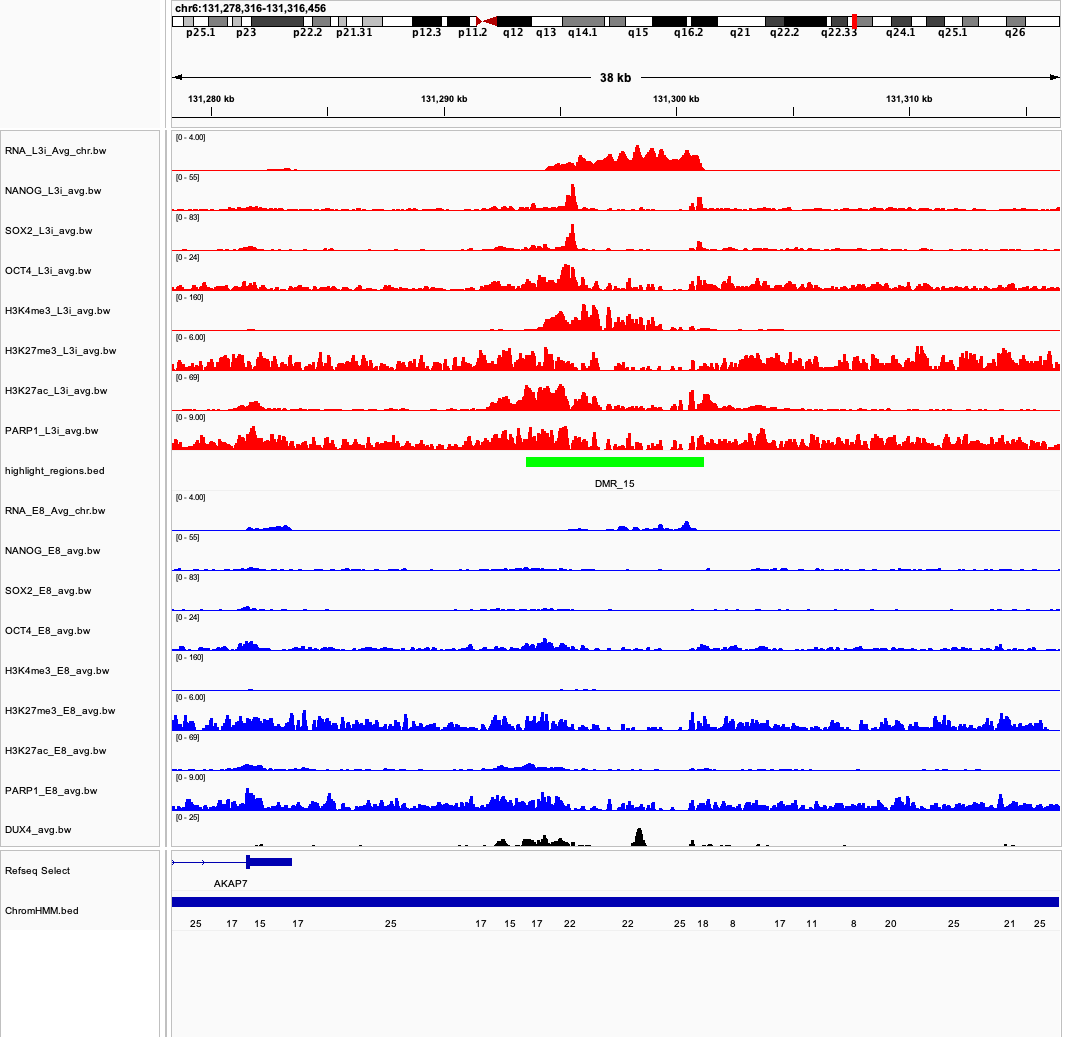
## Core transcription factors mediated the CpG methylation regulation of transcription by activating putative enhancer and super-enhancer.

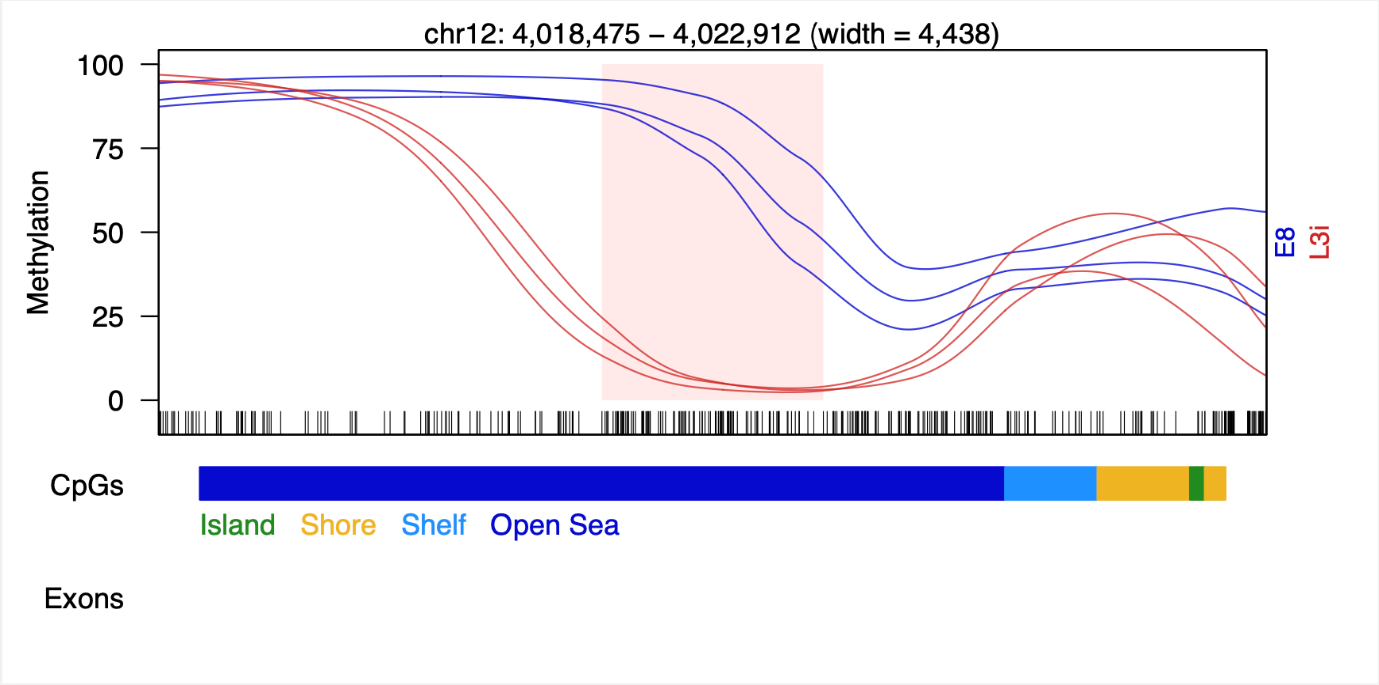
Although previous results suggested no direct regulation of transcription by methylation, it was still obvious that the local demethylation of TIRN-SC coincided with their overexpression of transcription factors in the undifferentiated state while no clear expression pattern could be found among transcription factors of TIRN-SC after retinal differentiation with reduced methylation differences from primed cells. Therefore, the transcriptional programs may be regulated by CpG methylation indirectly, putatively via the interactions with core transcription factor binding given the ChIP-seq results. An integrated analysis combining all sequencing results was performed on top DMRs with largest differences between TIRN-SC and the primed cells. Interestingly, a similar pattern was found in multiple DMRs, where strong transcription signals were reported at intergenic regions with higher H3K27ac levels that were annotated as or close to enhancer regions (**Fig. 28**). These non-coding RNAs (ncRNA) transcribed from enhancer regions could be putative enhancer RNAs (eRNA) and super enhancer RNAs (seRNA) based on the region lengths, which were reported to play essential roles in defining cell identity and regulating pluripotency especially when bound by core transcription factors55,56. Given that it was recently revealed that SON co-binding regions gained by TIRN-SC in the undifferentiated state were enriched for H3K27ac marked enhancers6, the indirect regulation of the blastomere-like transcriptional programs of TIRN-SC by CpG methylation would likely be mediated by core transcription factor binding that could activate putative eRNAs and seRNAs.

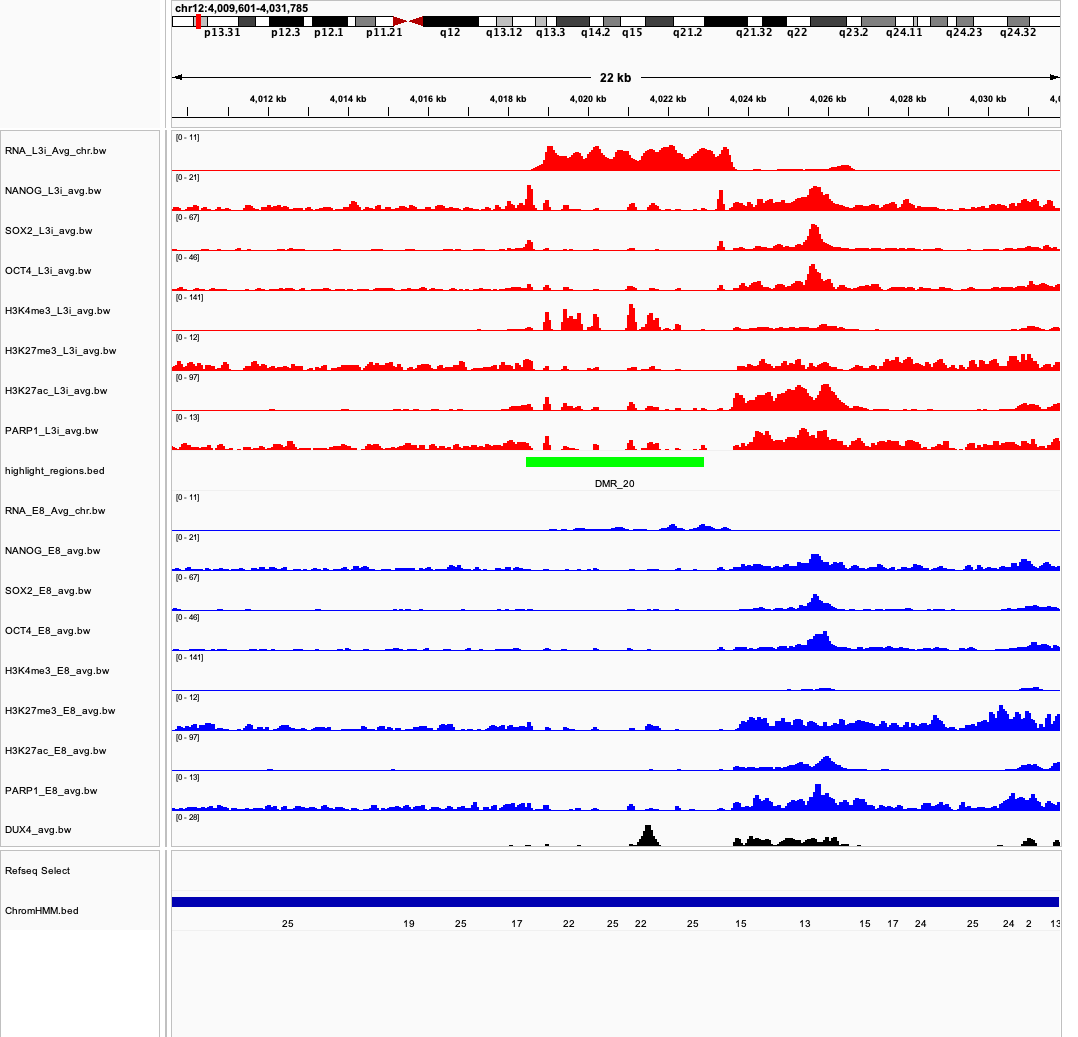












**Figure 28.**Sample IGV snapshots of RNA-Seq and ChIP-Seq score tracks for corresponding DMR in the undifferentiated state.

RNA-Seq tracks are the averaged CPM-normalized scores (n = 3) while ChIP-Seq tracks are the averaged RPGC-normalized scores (n = 2) for NANOG, SOX2, OCT4, SON co-binding, H3K4me3, H3K27me3, H3K27ac, and PARP1, with TIRN-SC in red and primed hPSC in blue. DUX4 ChIP-Seq signal is sourced from literature and therefore labeled in black. The DMR site labeled in pink rectangle in the DMR curve is highlighted by the green track in IGV snapshot. The ChromHMM numerically labeled states can be found in **Table 3**.

**Table 3.** ChromHMM 25-state model

|  |  |  |
| --- | --- | --- |
| **STATE NO.** | **MNEMONIC** | **DESCRIPTION** |
| 1 | TssA | Active TSS |
| 2 | PromU | Promoter Upstream TSS |
| 3 | PromD1 | Promoter Downstream TSS 1 |
| 4 | PromD2 | Promoter Downstream TSS 2 |
| 5 | Tx5 | Transcribed - 5' preferential |
| 6 | Tx | Strong transcription |
| 7 | Tx3 | Transcribed - 3' preferential |
| 8 | TxWk | Weak transcription |
| 9 | TxReg | Transcribed & regulatory (Prom/Enh) |
| 10 | TxEnh5 | Transcribed 5' preferential and Enh |
| 11 | TxEnh3 | Transcribed 3' preferential and Enh |
| 12 | TxEnhW | Transcribed and Weak Enhancer |
| 13 | EnhA1 | Active Enhancer 1 |
| 14 | EnhA2 | Active Enhancer 2 |
| 15 | EnhAF | Active Enhancer Flank |
| 16 | EnhW1 | Weak Enhancer 1 |
| 17 | EnhW2 | Weak Enhancer 2 |
| 18 | EnhAc | Primary H3K27ac possible Enhancer |
| 19 | DNase | Primary DNase |
| 20 | ZNF/Rpts | ZNF genes & repeats |
| 21 | Het | Heterochromatin |
| 22 | PromP | Poised Promoter |
| 23 | PromBiv | Bivalent Promoter |
| 24 | ReprPC | Repressed Polycomb |
| 25 | Quies | Quiescent/Low |

# Discussion

Naïve human PSC lines are pivotal models for developmental study at early embryonic stages, while disputes about their CpG methylation nature still exist in current acknowledgement8–11. Here, we present an integrative analysis of CpG methylation combined with multi-omics data in tankyrase/PARP inhibitor-regulated naïve stem cells (TIRN-SCs), which has been shown to exhibit transcriptomic similarity to early embryonic blastomeres and possess enhanced developmental plasticity6. The current study provides additional insights into the mechanisms of interplay between epigenetic dynamics and their complicated ultimate regulations on the transcriptional features enabling TIRN-SC blastomere-like identity and differentiation trajectory, with focus on how CpG methylation coordinates with protein-DNA interactions to activate key factor expression to establish and maintain this unique developmental state.

Contrary to the expected global hypomethylation repeatedly reported in naïve pluripotent stem cell models8,9, our TIRN-SC model exhibited only modest global methylation decrease. However, significant methylation reprogramming did occur in the undifferentiated state, with local demethylation observed in TIRN-SC at more than 10,000 differentially methylated regions (DMRs) that were highly enriched in regulatory regions, including CpG shores and promoters. In addition, differential binding sites of the core pluripotency transcription factors SOX2, OCT4, and NANOG (SON) experienced notable demethylation, which suggests a region-specific local methylation pattern restructuring chromatin accessibility and gene expression cooperated with SON regulation that all together build the primary foundation of blastomere-like identity in TIRN-SCs. Importantly, such methylation changes centered on SON binding sites coincide with and provide insights into the earlier proteogenomic findings from the Zambidis lab, which identified TIRN-SCs as the first model of global PARP inhibition-induced dedifferentiation of hPSC6. According to that groundbreaking advancement, the coordinated and interconnected downregulation of PARP1 and upregulation of TNKS1/2 led to a proteogenomic signatures mimicking the 2C–8C human embryos. PARP1 is proposed to play a decisive role in regulating genome-wide SON binding features by its direct interaction with OCT4 and SOX2 and mediating effect, such as histone modification and ubiquitin-proteasome system. It is likely that the decreased protein level of PARP1 in TIRN-SCs allows the de-repression of early developmental programs by enhancing SON binding and destabilizing transcriptional silencing complexes. These findings support a synergistic model in which PARP1 loss and SON redistribution cooperatively shape the transcriptional plasticity of the blastomere-like state.

Given that SON, the key components of such regulatory network, reportedly show co-binding features in enhancer regions55,56, our methylation analysis explored the potential mechanisms relating methylation to enhancers. We revealed that many top TIRN-SC hypomethylated DMRs in the undifferentiated state were distributed on or near intergenic regions exhibiting strong transcriptional activity, indicating possible non-coding RNAs. And the regulatory epigenetic modifications on these regions, especially the higher levels of H3K27ac, further confirmed their identities of putative enhancer RNAs (eRNAs) or super-enhancer RNAs (seRNAs) based on lengths55,57,58. Their transcription typically means an increased local chromatin accessibility and attributes such phenomenon to SON binding and local demethylation in this case. These non-coding RNAs have been shown to recruit transcriptional complexes and stabilize enhancer-promoter interactions 59,60. Thus, CpG methylation may not only drive binding of core transcription factors but also further base on that to activate enhancer-associated transcription, reinforcing the regulatory circuitry contributing to the blastomere-like phenotype.

Interestingly, although the SON binding and enhancer activation have been proved to have significant associations with CpG methylation, our analysis showed that the downstream gene expression could hardly be correlated with DNA methylation, which suggests that CpG methylation does not directly regulate gene expression. Instead, it may reprogram the transcriptomics potentially via interactions with other mediating factors. Such an indirect regulatory mechanism corresponds with our observations that demethylation events can coincide with transcription factor binding sites and regulatory elements, rather than gene bodies or promoters alone. After differentiation into retinal organoids, the differential methylation and transcription of TIRN-SC compared to primed cells largely diminished, meaning that their blastomere-like state is not irreversible due to fixed demethylation but can still be differentiated into another developmental state by epigenetic and proteogenomic reprogramming. While it has been reported that TIRN-SC-derived ROs can achieve improved morphological features15,16, GSEA shows suppression of retinal and neural development pathways in them. Such inconsistence suggests that phenotypic maturity may not be completely explained by sole transcriptional signatures and post-translational modifications need to be further investigated.

All these findings above improve our understanding of the TIRN-SC model in which CpG methylation functions as a key component in the PARP1/TNKS-mediated overall proteogenomic mechanisms behind the blastomere-like pluripotency in addition to transcription factors and enhancers. In the absence of significant global remodeling, TIRN-SC reprogramming appears to be governed by local changes at key regulatory loci with modifications of methylation, transcription factor binding, ubiquitination, chromatin accessibility, and eventually transcription all combined. In the context of regenerative medicine, this work highlights the importance of fine-tuning local methylation in stem cell derived tissue development and transplantation. Future studies should explore the potential of improving the stability and function of engineered stem cell states after differentiation via targeted and accurate methylation modifications.

Ultimately, our multi-omics analysis underscores the importance of integrating methylation, chromatin accessibility, transcription, protein-DNA interaction and proteomic data to understand the mechanisms of stem cell regulation and pluripotency dynamics. As previously described and reinforced here, the functional human blastomere-like TIRN-SCs provide an invaluable model for understanding early human embryonic development and enhancing the utility of pluripotent cells in clinical applications.

# Limitations

The current study verified the previously described functional blastomere-like state of TIRN-SC and further suggested the key regulatory role of local methylation in transcriptional reprogramming through interplay with core transcription factor binding and enhancer activation. However, sequencing data evidence was not strictly rigorous, with replicates consisting of single samples from three separate cell lines and inconsistent sampling between different sequencing studies (only RUES02 for WGSB) despite previous statement of the elimination of interline variability61. Similarly, although the diminished epigenetic signature differences between TIRN-SC and primed cells were revealed here, the data inconsistency issue also existed (no hiPSC (E5C3) comparison available). In addition, the opposing transcriptomic (suppressed phototransduction pathway) and physiology (improved maturation and survival) reflected the complexity of the retinal differentiation considering the complicated nature of retinal structures, suggesting critical need for single-cell analysis. To further explore the epigenetic dynamics of TIRN-SC, future studied may also be benefited by introducing time-course datasets to increase temporal resolution and by including other differentiation trajectories to improve comprehensiveness.

# Bibliography

(1) Thomson, J. A.; Itskovitz-Eldor, J.; Shapiro, S. S.; Waknitz, M. A.; Swiergiel, J. J.; Marshall, V. S.; Jones, J. M. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* **1998**, *282* (5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145.

(2) Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **2007**, *131* (5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019.

(3) Nichols, J.; Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* **2009**, *4* (6), 487–492. https://doi.org/10.1016/j.stem.2009.05.015.

(4) Weinberger, L.; Ayyash, M.; Novershtern, N.; Hanna, J. H. Dynamic Stem Cell States: Naive to Primed Pluripotency in Rodents and Humans. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (3), 155–169. https://doi.org/10.1038/nrm.2015.28.

(5) Guo, G.; von Meyenn, F.; Rostovskaya, M.; Clarke, J.; Dietmann, S.; Baker, D.; Sahakyan, A.; Myers, S.; Bertone, P.; Reik, W.; Plath, K.; Smith, A. Epigenetic Resetting of Human Pluripotency. *Dev. Camb. Engl.* **2017**, *144* (15), 2748–2763. https://doi.org/10.1242/dev.146811.

(6) Zimmerlin, L.; Angarita, A.; Park, T. S.; Evans-Moses, R.; Thomas, J.; Yan, S.; Uribe, I.; Vegas, I.; Kochendoerfer, C.; Buys, W.; Leung, A. K. L.; Zambidis, E. T. Proteogenomic Reprogramming to a Functional Human Totipotent Stem Cell State via a PARP-DUX4 Regulatory Axis. bioRxiv June 15, 2024, p 2024.06.14.598510. https://doi.org/10.1101/2024.06.14.598510.

(7) Bird, A. DNA Methylation Patterns and Epigenetic Memory. *Genes Dev.* **2002**, *16* (1), 6–21. https://doi.org/10.1101/gad.947102.

(8) Pastor, W. A.; Chen, D.; Liu, W.; Kim, R.; Sahakyan, A.; Lukianchikov, A.; Plath, K.; Jacobsen, S. E.; Clark, A. T. Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* **2016**, *18* (3), 323–329. https://doi.org/10.1016/j.stem.2016.01.019.

(9) Theunissen, T. W.; Jaenisch, R. Mechanisms of Gene Regulation in Human Embryos and Pluripotent Stem Cells. *Dev. Camb. Engl.* **2017**, *144* (24), 4496–4509. https://doi.org/10.1242/dev.157404.

(10) Lee, H. J.; Hore, T. A.; Reik, W. Reprogramming the Methylome: Erasing Memory and Creating Diversity. *Cell Stem Cell* **2014**, *14* (6), 710–719. https://doi.org/10.1016/j.stem.2014.05.008.

(11) Messerschmidt, D. M.; Knowles, B. B.; Solter, D. DNA Methylation Dynamics during Epigenetic Reprogramming in the Germline and Preimplantation Embryos. *Genes Dev.* **2014**, *28* (8), 812–828. https://doi.org/10.1101/gad.234294.113.

(12) Stadler, M. B.; Murr, R.; Burger, L.; Ivanek, R.; Lienert, F.; Schöler, A.; Nimwegen, E. van; Wirbelauer, C.; Oakeley, E. J.; Gaidatzis, D.; Tiwari, V. K.; Schübeler, D. DNA-Binding Factors Shape the Mouse Methylome at Distal Regulatory Regions. *Nature* **2011**, *480* (7378), 490–495. https://doi.org/10.1038/nature10716.

(13) Soufi, A.; Garcia, M. F.; Jaroszewicz, A.; Osman, N.; Pellegrini, M.; Zaret, K. S. Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. *Cell* **2015**, *161* (3), 555–568. https://doi.org/10.1016/j.cell.2015.03.017.

(14) Kilens, S.; Meistermann, D.; Moreno, D.; Chariau, C.; Gaignerie, A.; Reignier, A.; Lelièvre, Y.; Casanova, M.; Vallot, C.; Nedellec, S.; Flippe, L.; Firmin, J.; Song, J.; Charpentier, E.; Lammers, J.; Donnart, A.; Marec, N.; Deb, W.; Bihouée, A.; Le Caignec, C.; Pecqueur, C.; Redon, R.; Barrière, P.; Bourdon, J.; Pasque, V.; Soumillon, M.; Mikkelsen, T. S.; Rougeulle, C.; Fréour, T.; David, L. Parallel Derivation of Isogenic Human Primed and Naive Induced Pluripotent Stem Cells. *Nat. Commun.* **2018**, *9* (1), 360. https://doi.org/10.1038/s41467-017-02107-w.

(15) Zambidis, E.; Bhutto, I. A.; Kanherkar, R. T.; Park, T. S.; Moses, R. T.; Zimmerlin, L. T.; Liu, Y. T.; Singh, M. S.; Lutty, G. A. Long-Term Engraftment of Mature Human Photoreceptors Following Subretinal Transplantation of Naïve hPSC-Derived Retinal Organoids. *Invest. Ophthalmol. Vis. Sci.* **2020**, *61* (7), 5203.

(16) Zimmerlin, L.; Bhutto, I. A.; Kanherkar, R.; Park, T. S.; Barbato, M.; Koldobskiy, M.; Liu, Y.; Singh, M.; Lutty, G. A.; Zambidis, E. T. Improved Generation and Long-Term Engraftment of Retinal Organoids from Tankyrase/PARP-Inhibitor-Regulated Naïve Human Pluripotent Stem Cells (TIRN-hPSC). *Invest. Ophthalmol. Vis. Sci.* **2022**, *63* (7), 3725-F0331.

(17) Zimmerlin, L.; Park, T. S.; Huo, J. S.; Verma, K.; Pather, S. R.; Talbot, C. C.; Agarwal, J.; Steppan, D.; Zhang, Y. W.; Considine, M.; Guo, H.; Zhong, X.; Gutierrez, C.; Cope, L.; Canto-Soler, M. V.; Friedman, A. D.; Baylin, S. B.; Zambidis, E. T. Tankyrase Inhibition Promotes a Stable Human Naïve Pluripotent State with Improved Functionality. *Dev. Camb. Engl.* **2016**, *143* (23), 4368–4380. https://doi.org/10.1242/dev.138982.

(18) Park, T. S.; Zimmerlin, L.; Evans-Moses, R.; Thomas, J.; Huo, J. S.; Kanherkar, R.; He, A.; Ruzgar, N.; Grebe, R.; Bhutto, I. Vascular Progenitors Generated from Tankyrase Inhibitor-Regulated Naïve Diabetic Human iPSC Potentiate Efficient Revascularization of Ischemic Retina. *Nat. Commun.* **2020**, *11* (1), 1195.

(19) Park, T. S.; Zimmerlin, L.; Evans-Moses, R.; Zambidis, E. T. Chemical Reversion of Conventional Human Pluripotent Stem Cells to a Naïve-like State with Improved Multilineage Differentiation Potency. *J. Vis. Exp. JoVE* **2018**, No. 136, 57921. https://doi.org/10.3791/57921.

(20) Meyer, J. S.; Shearer, R. L.; Capowski, E. E.; Wright, L. S.; Wallace, K. A.; McMillan, E. L.; Zhang, S.-C.; Gamm, D. M. Modeling Early Retinal Development with Human Embryonic and Induced Pluripotent Stem Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (39), 16698–16703. https://doi.org/10.1073/pnas.0905245106.

(21) Meyer, J. S.; Howden, S. E.; Wallace, K. A.; Verhoeven, A. D.; Wright, L. S.; Capowski, E. E.; Pinilla, I.; Martin, J. M.; Tian, S.; Stewart, R.; Pattnaik, B.; Thomson, J. A.; Gamm, D. M. Optic Vesicle-like Structures Derived from Human Pluripotent Stem Cells Facilitate a Customized Approach to Retinal Disease Treatment. *Stem Cells Dayt. Ohio* **2011**, *29* (8), 1206–1218. https://doi.org/10.1002/stem.674.

(22) Phillips, M. J.; Wallace, K. A.; Dickerson, S. J.; Miller, M. J.; Verhoeven, A. D.; Martin, J. M.; Wright, L. S.; Shen, W.; Capowski, E. E.; Percin, E. F.; Perez, E. T.; Zhong, X.; Canto-Soler, M. V.; Gamm, D. M. Blood-Derived Human iPS Cells Generate Optic Vesicle–Like Structures with the Capacity to Form Retinal Laminae and Develop Synapses. *Invest. Ophthalmol. Vis. Sci.* **2012**, *53* (4), 2007–2019. https://doi.org/10.1167/iovs.11-9313.

(23) Zhong, X.; Gutierrez, C.; Xue, T.; Hampton, C.; Vergara, M. N.; Cao, L.-H.; Peters, A.; Park, T. S.; Zambidis, E. T.; Meyer, J. S.; Gamm, D. M.; Yau, K.-W.; Canto-Soler, M. V. Generation of Three-Dimensional Retinal Tissue with Functional Photoreceptors from Human iPSCs. *Nat. Commun.* **2014**, *5* (1), 4047. https://doi.org/10.1038/ncomms5047.

(24) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.

(25) Ramírez, F.; Ryan, D. P.; Grüning, B.; Bhardwaj, V.; Kilpert, F.; Richter, A. S.; Heyne, S.; Dündar, F.; Manke, T. deepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis. *Nucleic Acids Res.* **2016**, *44* (W1), W160-165. https://doi.org/10.1093/nar/gkw257.

(26) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (12), 550. https://doi.org/10.1186/s13059-014-0550-8.

(27) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43* (7), e47. https://doi.org/10.1093/nar/gkv007.

(28) Gu, Z.; Eils, R.; Schlesner, M. Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data. *Bioinformatics* **2016**, *32* (18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

(29) Zou, Z.; Zhang, C.; Wang, Q.; Hou, Z.; Xiong, Z.; Kong, F.; Wang, Q.; Song, J.; Liu, B.; Liu, B.; Wang, L.; Lai, F.; Fan, Q.; Tao, W.; Zhao, S.; Ma, X.; Li, M.; Wu, K.; Zhao, H.; Chen, Z.-J.; Xie, W. Translatome and Transcriptome Co-Profiling Reveals a Role of TPRXs in Human Zygotic Genome Activation. *Science* **2022**, *378* (6615), abo7923. https://doi.org/10.1126/science.abo7923.

(30) Hoshino, A.; Ratnapriya, R.; Brooks, M. J.; Chaitankar, V.; Wilken, M. S.; Zhang, C.; Starostik, M. R.; Gieser, L.; La Torre, A.; Nishio, M.; Bates, O.; Walton, A.; Bermingham-McDonogh, O.; Glass, I. A.; Wong, R. O. L.; Swaroop, A.; Reh, T. A. Molecular Anatomy of the Developing Human Retina. *Dev. Cell* **2017**, *43* (6), 763-779.e4. https://doi.org/10.1016/j.devcel.2017.10.029.

(31) Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS J. Integr. Biol.* **2012**, *16* (5), 284–287. https://doi.org/10.1089/omi.2011.0118.

(32) Croft, D.; O’Kelly, G.; Wu, G.; Haw, R.; Gillespie, M.; Matthews, L.; Caudy, M.; Garapati, P.; Gopinath, G.; Jassal, B.; Jupe, S.; Kalatskaya, I.; Mahajan, S.; May, B.; Ndegwa, N.; Schmidt, E.; Shamovsky, V.; Yung, C.; Birney, E.; Hermjakob, H.; D’Eustachio, P.; Stein, L. Reactome: A Database of Reactions, Pathways and Biological Processes. *Nucleic Acids Res.* **2011**, *39* (Database issue), D691–D697. https://doi.org/10.1093/nar/gkq1018.

(33) Danecek, P.; Bonfield, J. K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M. O.; Whitwham, A.; Keane, T.; McCarthy, S. A.; Davies, R. M.; Li, H. Twelve Years of SAMtools and BCFtools. *GigaScience* **2021**, *10* (2), giab008. https://doi.org/10.1093/gigascience/giab008.

(34) Li, Q.; Brown, J. B.; Huang, H.; Bickel, P. J. Measuring Reproducibility of High-Throughput Experiments. *Ann. Appl. Stat.* **2011**, *5* (3), 1752–1779. https://doi.org/10.1214/11-AOAS466.

(35) Quinlan, A. R.; Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26* (6), 841–842. https://doi.org/10.1093/bioinformatics/btq033.

(36) Hendrickson, P. G.; Doráis, J. A.; Grow, E. J.; Whiddon, J. L.; Lim, J.-W.; Wike, C. L.; Weaver, B. D.; Pflueger, C.; Emery, B. R.; Wilcox, A. L.; Nix, D. A.; Peterson, C. M.; Tapscott, S. J.; Carrell, D. T.; Cairns, B. R. Conserved Roles of Mouse DUX and Human DUX4 in Activating Cleavage-Stage Genes and MERVL/HERVL Retrotransposons. *Nat. Genet.* **2017**, *49* (6), 925–934. https://doi.org/10.1038/ng.3844.

(37) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4), 357–359. https://doi.org/10.1038/nmeth.1923.

(38) Zhang, Y.; Liu, T.; Meyer, C. A.; Eeckhoute, J.; Johnson, D. S.; Bernstein, B. E.; Nusbaum, C.; Myers, R. M.; Brown, M.; Li, W.; Liu, X. S. Model-Based Analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9* (9), R137. https://doi.org/10.1186/gb-2008-9-9-r137.

(39) Yu, G.; Wang, L.-G.; He, Q.-Y. ChIPseeker: An R/Bioconductor Package for ChIP Peak Annotation, Comparison and Visualization. *Bioinformatics* **2015**, *31* (14), 2382–2383. https://doi.org/10.1093/bioinformatics/btv145.

(40) Krueger, F. Trim Galore. *Wrapper Tool Cutadapt FastQC Consistently Apply Qual. Adapt. Trimming FastQ Files* **2015**, *516* (517), 517.

(41) Krueger, F.; Andrews, S. R. Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *bioinformatics* **2011**, *27* (11), 1571–1572.

(42) Laufer, B. I.; Hwang, H.; Jianu, J. M.; Mordaunt, C. E.; Korf, I. F.; Hertz-Picciotto, I.; LaSalle, J. M. Low-Pass Whole Genome Bisulfite Sequencing of Neonatal Dried Blood Spots Identifies a Role for RUNX1 in Down Syndrome DNA Methylation Profiles. *Hum. Mol. Genet.* **2021**, *29* (21), 3465–3476. https://doi.org/10.1093/hmg/ddaa218.

(43) Wickham, H. Introduction. In *ggplot2: Elegant Graphics for Data Analysis*; Wickham, H., Ed.; Springer International Publishing: Cham, 2016; pp 3–10. https://doi.org/10.1007/978-3-319-24277-4\_1.

(44) Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P. Integrative Genomics Viewer. *Nat. Biotechnol.* **2011**, *29* (1), 24–26. https://doi.org/10.1038/nbt.1754.

(45) Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; Ziller, M. J.; Amin, V.; Whitaker, J. W.; Schultz, M. D.; Ward, L. D.; Sarkar, A.; Quon, G.; Sandstrom, R. S.; Eaton, M. L.; Wu, Y.-C.; Pfenning, A. R.; Wang, X.; Claussnitzer, M.; Yaping Liu; Coarfa, C.; Alan Harris, R.; Shoresh, N.; Epstein, C. B.; Gjoneska, E.; Leung, D.; Xie, W.; David Hawkins, R.; Lister, R.; Hong, C.; Gascard, P.; Mungall, A. J.; Moore, R.; Chuah, E.; Tam, A.; Canfield, T. K.; Scott Hansen, R.; Kaul, R.; Sabo, P. J.; Bansal, M. S.; Carles, A.; Dixon, J. R.; Farh, K.-H.; Feizi, S.; Karlic, R.; Kim, A.-R.; Kulkarni, A.; Li, D.; Lowdon, R.; Elliott, G.; Mercer, T. R.; Neph, S. J.; Onuchic, V.; Polak, P.; Rajagopal, N.; Ray, P.; Sallari, R. C.; Siebenthall, K. T.; Sinnott-Armstrong, N. A.; Stevens, M.; Thurman, R. E.; Wu, J.; Zhang, B.; Zhou, X.; Beaudet, A. E.; Boyer, L. A.; De Jager, P. L.; Farnham, P. J.; Fisher, S. J.; Haussler, D.; Jones, S. J. M.; Li, W.; Marra, M. A.; McManus, M. T.; Sunyaev, S.; Thomson, J. A.; Tlsty, T. D.; Tsai, L.-H.; Wang, W.; Waterland, R. A.; Zhang, M. Q.; Chadwick, L. H.; Bernstein, B. E.; Costello, J. F.; Ecker, J. R.; Hirst, M.; Meissner, A.; Milosavljevic, A.; Ren, B.; Stamatoyannopoulos, J. A.; Wang, T.; Kellis, M. Integrative Analysis of 111 Reference Human Epigenomes. *Nature* **2015**, *518* (7539), 317–330. https://doi.org/10.1038/nature14248.

(46) Ziller, M. J.; Gu, H.; Müller, F.; Donaghey, J.; Tsai, L. T.-Y.; Kohlbacher, O.; De Jager, P. L.; Rosen, E. D.; Bennett, D. A.; Bernstein, B. E.; Gnirke, A.; Meissner, A. Charting a Dynamic DNA Methylation Landscape of the Human Genome. *Nature* **2013**, *500* (7463), 477–481. https://doi.org/10.1038/nature12433.

(47) Doi, A.; Park, I.-H.; Wen, B.; Murakami, P.; Aryee, M. J.; Irizarry, R.; Herb, B.; Ladd-Acosta, C.; Rho, J.; Loewer, S.; Miller, J.; Schlaeger, T.; Daley, G. Q.; Feinberg, A. P. Differential Methylation of Tissue- and Cancer-Specific CpG Island Shores Distinguishes Human Induced Pluripotent Stem Cells, Embryonic Stem Cells and Fibroblasts. *Nat. Genet.* **2009**, *41* (12), 1350–1353. https://doi.org/10.1038/ng.471.

(48) Waite, J. B.; Boytz, R.; Traeger, A. R.; Lind, T. M.; Lumbao-Conradson, K.; Torigoe, S. E. A Suboptimal OCT4-SOX2 Binding Site Facilitates the Naïve-State Specific Function of a Klf4 Enhancer. *PloS One* **2024**, *19* (9), e0311120. https://doi.org/10.1371/journal.pone.0311120.

(49) Huang, X.; Park, K.-M.; Gontarz, P.; Zhang, B.; Pan, J.; McKenzie, Z.; Fischer, L. A.; Dong, C.; Dietmann, S.; Xing, X.; Shliaha, P. V.; Yang, J.; Li, D.; Ding, J.; Lungjangwa, T.; Mitalipova, M.; Khan, S. A.; Imsoonthornruksa, S.; Jensen, N.; Wang, T.; Kadoch, C.; Jaenisch, R.; Wang, J.; Theunissen, T. W. OCT4 Cooperates with Distinct ATP-Dependent Chromatin Remodelers in Naïve and Primed Pluripotent States in Human. *Nat. Commun.* **2021**, *12* (1), 5123. https://doi.org/10.1038/s41467-021-25107-3.

(50) Buecker, C.; Srinivasan, R.; Wu, Z.; Calo, E.; Acampora, D.; Faial, T.; Simeone, A.; Tan, M.; Swigut, T.; Wysocka, J. Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. *Cell Stem Cell* **2014**, *14* (6), 838–853. https://doi.org/10.1016/j.stem.2014.04.003.

(51) Jones, M. K.; Agarwal, D.; Mazo, K. W.; Chopra, M.; Jurlina, S. L.; Dash, N.; Xu, Q.; Ogata, A. R.; Chow, M.; Hill, A. D.; Kambli, N. K.; Xu, G.; Sasik, R.; Birmingham, A.; Fisch, K. M.; Weinreb, R. N.; Enke, R. A.; Skowronska-Krawczyk, D.; Wahlin, K. J. Chromatin Accessibility and Transcriptional Differences in Human Stem Cell-Derived Early-Stage Retinal Organoids. *Cells* **2022**, *11* (21), 3412. https://doi.org/10.3390/cells11213412.

(52) Abdolvand, N.; Tostoes, R.; Raimes, W.; Kumar, V.; Szita, N.; Veraitch, F. Long-Term Retinal Differentiation of Human Induced Pluripotent Stem Cells in a Continuously Perfused Microfluidic Culture Device. *Biotechnol. J.* **2019**, *14* (3), e1800323. https://doi.org/10.1002/biot.201800323.

(53) Petropoulos, S.; Edsgärd, D.; Reinius, B.; Deng, Q.; Panula, S. P.; Codeluppi, S.; Plaza Reyes, A.; Linnarsson, S.; Sandberg, R.; Lanner, F. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **2016**, *165* (4), 1012–1026. https://doi.org/10.1016/j.cell.2016.03.023.

(54) Yan, L.; Yang, M.; Guo, H.; Yang, L.; Wu, J.; Li, R.; Liu, P.; Lian, Y.; Zheng, X.; Yan, J.; Huang, J.; Li, M.; Wu, X.; Wen, L.; Lao, K.; Li, R.; Qiao, J.; Tang, F. Single-Cell RNA-Seq Profiling of Human Preimplantation Embryos and Embryonic Stem Cells. *Nat. Struct. Mol. Biol.* **2013**, *20* (9), 1131–1139. https://doi.org/10.1038/nsmb.2660.

(55) Whyte, W. A.; Orlando, D. A.; Hnisz, D.; Abraham, B. J.; Lin, C. Y.; Kagey, M. H.; Rahl, P. B.; Lee, T. I.; Young, R. A. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **2013**, *153* (2), 307–319. https://doi.org/10.1016/j.cell.2013.03.035.

(56) Ma, H.; Qu, J.; Pang, Z.; Luo, J.; Yan, M.; Xu, W.; Zhuang, H.; Liu, L.; Qu, Q. Super-Enhancer Omics in Stem Cell. *Mol. Cancer* **2024**, *23* (1), 153. https://doi.org/10.1186/s12943-024-02066-z.

(57) Creyghton, M. P.; Cheng, A. W.; Welstead, G. G.; Kooistra, T.; Carey, B. W.; Steine, E. J.; Hanna, J.; Lodato, M. A.; Frampton, G. M.; Sharp, P. A.; Boyer, L. A.; Young, R. A.; Jaenisch, R. Histone H3K27ac Separates Active from Poised Enhancers and Predicts Developmental State. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (50), 21931–21936. https://doi.org/10.1073/pnas.1016071107.

(58) Hnisz, D.; Abraham, B. J.; Lee, T. I.; Lau, A.; Saint-André, V.; Sigova, A. A.; Hoke, H. A.; Young, R. A. Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **2013**, *155* (4), 934–947. https://doi.org/10.1016/j.cell.2013.09.053.

(59) Li, W.; Notani, D.; Rosenfeld, M. G. Enhancers as Non-Coding RNA Transcription Units: Recent Insights and Future Perspectives. *Nat. Rev. Genet.* **2016**, *17* (4), 207–223. https://doi.org/10.1038/nrg.2016.4.

(60) Lando, D.; Ma, X.; Cao, Y.; Jartseva, A.; Stevens, T. J.; Boucher, W.; Reynolds, N.; Montibus, B.; Hall, D.; Lackner, A.; Ragheb, R.; Leeb, M.; Hendrich, B. D.; Laue, E. D. Enhancer-Promoter Interactions Are Reconfigured through the Formation of Long-Range Multiway Hubs as Mouse ES Cells Exit Pluripotency. *Mol. Cell* **2024**, *84* (8), 1406-1421.e8. https://doi.org/10.1016/j.molcel.2024.02.015.

(61) Zimmerlin, L.; Park, T. S.; Bhutto, I.; Lutty, G.; Zambidis, E. T. Generation of Pericytic-Vascular ProgenitorsPericytic-Vascular Progenitors from Tankyrase/PARP-Inhibitor-Regulated Naïve (TIRN)Tankyrase/PARP-Inhibitor-Regulated Naïve (TIRN) Human Pluripotent Stem Cells. In *Human Naïve Pluripotent Stem Cells*; Rugg-Gunn, P., Ed.; Springer US: New York, NY, 2022; pp 133–156. https://doi.org/10.1007/978-1-0716-1908-7\_10.