CHAPTER 3

Trans-omics analysis of neural cell fates using stem cell-derived brain organoids

---

This work is published in the following journal:

Chen, C., Lee, S., Zyner, KG., Fernando, M., Nemeruck, V., Wong, E., Marshall, LL., Wark, JW., Arymanesh, N., Tam, PPL., Graham, ME., Gonzalez-Cordero, A., Yang, P. (2024) Trans-omic Profiling Uncovers Molecular Controls of the Early Human Cerebral Organoid Formation. *Cell Reports*, 43, 5. 10.1016/j.celrep.2024.114219

My contribution: I pre-processed and analysed all the multi-omic data and additional experimental data generated in this study. I generated all the visualisations and interpreted the data accordingly. I also led the design of the additional validation experiments and scRNA-seq.

3.1 Motivation and overview

While spatial transcriptomics provides contextualisation to gene expression in cells, biological systems are inherently complex and involve multiple layers of regulation beyond the transcriptome. In this chapter, I apply this philosophy to the integrative analysis of multiple omics data to reconstruct the signaling and gene regulatory networks (GRNs) that drive neural lineage commitment in stem-cell derived brain organoids, offering deeper insights into the molecular mechanisms of early neurogenesis.

3D organoid cultures provide the flexibility to model developmental processes without the need to acquire *in vivo* samples. Particularly in the case of human development, there are ethical constraints with obtaining tissue at such an early stage of life. Many existing studies have profiled the single-cell transcriptome, bulk transcriptome or epigenome of human brain organoid development starting from complete neural lineage commitment to maturity. To complement the studies in the field using organoid models as a proxy to study human brain development, I focused on characterising the myriad of differentiation and lineage commitment signals that are made from the immediate exit of pluripotency to the acquisition of early neural fate for the directed outcome of optimising production of cerebral organoids. Here, we generated the phosphoproteomic, proteomic, and (single-cell)-transcriptomic profiles of early brain organoid development at matching timepoints in parallel to understand the signalling networks initiated by the phosphorylation of key lineage-driving molecular factors over time. I explored how the phospho-signalling further propagates to downstream transcriptional and translational changes which shape neural fate in the cerebral organoid and show the transcriptional fidelity by demonstrating the accurate spatial localisation of cell identity network activity to *in vivo* brain tissue. Together, this is the first organoid study that integratively analyses phosphoproteomics with transcriptomics and proteomics to understand key factors in early neural lineage specification.

3.2 Introduction

Neural induction and human brain development are highly regulated processes orchestrated by various molecular and cellular programs that cut across cell signalling and gene transcription. Due to the limited access to human embryonic and fetal brain tissues, neurodevelopment was primarily interrogated through post-mortem dissections of brain tissue and the use of mouse models. However, the confounders of post-mortem samples and the discrepancies between mouse models and their human counterpart have posed significant challenges for understanding the molecular programs that govern this fundamental and highly dynamic process of human development [120]. Innovations in biotechnologies have allowed us to circumvent these issues with the use of human pluripotent stem cells (hPSCs) [121] for generating self-assembling 3D *in vitro* neural structures, so-called ‘human cerebral organoids (hCOs)’ [122, 55, 56, 123]. The capability of hCOs to recapitulate the self-organisation, structure and microenvironment of the human brain have made them an appealing model for studying human brain organogenesis and diseases [124, 125, 126].

Significant advances have been made in understanding the temporal progression and molecular pathways involved in neural differentiation in mammalian brain development [127, 128]. In particular, many studies have profiled the global transcriptional changes [129, 130, 131] to identify important genes that drive the neural differentiation process. Yet, it is well recognised that complex coordination of upstream signalling cascades with downstream GRNs is required for proper neurogenesis. The advancement of large-scale multi-omics profiling at the population- and single cell-level offers an exciting opportunity to characterise the key drivers across multiple molecular programs during brain development [132]. To this end, recent studies have profiled *in vitro* hCOs to elucidate the transcriptional and proteomic regulation occurring during neurodevelopment [132, 61, 62, 60, 57]. Nevertheless, cell signalling cascades and their interaction with downstream transcriptional regulation, together referred to as 'trans-omics' [22], are critical in controlling neurodevelopment [133] and have remained largely unexplored.

To gain insights into the global cell signalling landscapes and their downstream regulation of GRNs during early human brain development, we generated temporal trans-omic data profiling global phosphoproteome, proteome, and transcriptome at matched timepoints from hPSC aggregates exiting pluripotency to the generation of the neuroepithelium during early hCO formation [55]. Integrative analysis enabled the identification of key phospho-signalling events and their crosstalk with downstream GRNs during the neural lineage differentiation process. Notably, we found that phospho-signalling preferentially converges on transcriptional (co-)factors and chromatin remodellers. Single-cell RNA-sequencing (scRNA-seq) linked key regulators to cell subpopulations, revealing the specification of heterogeneous cell types in early hCOs. Assessment of the fidelity using *in vivo* data demonstrated that our *in vitro* hCO differentiation data recapitulated those of early human and mouse brains profiled using single-cell and spatial transcriptomics technologies. Finally, we experimentally evaluated SC79, a small molecule that activates AKT signalling, for controlling the hCO differentiation process, providing a guide for future development and optimisation of their differentiation protocols. We anticipate that this trans-omic resource of early hCO formation will provide a roadmap of neurogenesis in humans and contribute towards a greater understanding of developmental neurobiology and neurodevelopmental disorders.

3.3 Materials and methods

3.3.1 Maintenance of human PSCs

H9 ESCs (hPSCs) were maintained in 6 well plates on Geltrex (Thermo Fisher Scientific, #A1413302) matrix with daily media changes using Essential 8 media (Thermo Fisher

Scientific, #A1517001). We found four and a half 6 well plates were sufficient to provide 33 low attachment 96 well plates of organoids. This enabled harvesting of sufficient material at each time point for the analysis described.

3.3.2 Differentiation of hPSCs into hCOs

Once the hPSCs wells reached 70-80% confluency, two wells were harvested using lysis buffer and a cell scraper and snap frozen. The remaining wells were dissociated to single cells with warm Accutase (Thermo Fisher Scientific, #A1110501) and a modified version of the whole brain differentiation protocol described in [55] was followed. After centrifugation at 200G, the cells were resuspended in 10mL of plating media which consisted of Essential 6 (E6) (Thermo Fisher Scientific, #A1516401) with 4 ng/mL bFGF (Peprotech, #100-18C) and 50μM Y-27632 dihydrochloride (STEMCELL Technologies, #72307). E6 was used at the start of the differentiation to minimise stress to the hPSCs that are normally cultured in Essential 8 media. hPSC aggregates formed in E6, as opposed to E8 medium described in Lancaster et al. [55], showed less cell death and fewer debris and appeared larger by day 2 confirming improved cell survival.

Cells were counted and the cell suspension diluted to provide 9000 cells per 150μL and added to 33 low attachment 96-well plates (U-bottom plate, Thermo Fisher Scientific, #174925) using a multichannel pipette. Day 2 of differentiation was determined to be when cells were added to the wells. The cells were observed 24 hours later to ensure uniform organoids had formed with clear borders. 15 minutes prior to the 48 hours time point after plating (day zero of neuralisation), organoids and media were quickly collected into a 50mL tube. The cell suspension was centrifuged at 200G and most of the media was aspirated, leaving about 1mL with the organoids. The media was removed from the organoids and snap frozen on dry ice and the organoids were resuspended in lysis buffer and lysed using an electrically driven pestle. Organoids were also collected 1 hour and 6 hours later and on days 0, 1, 2, 4, 6 and 8 of differentiation.

For the remaining plates, half the media was removed and replaced with 150μL of Proneural Induction Media (PIM, Advanced DMEM, Thermo Fisher Scientific, #12634028; 1x N2, Thermo Fisher Scientific, # 17502048; 1x MEM NEAA,Thermo Fisher Scientific, #11140050; 2mM L-Glutamine, Thermo Fisher Scientific, #25030081; 1x Penicillin/Streptomycin, Thermo Fisher Scientific, #15070063; 0.5 μg Amphotericin B, Thermo Fisher

Scientific, #15290018). After another 48 hours, each organoid was transferred to a well of a low attachment 24 well plate containing 500µL of PIM media. Half media change with PIM was continued every other day for 4 days and on day 6 of differentiation the media was completely replaced with CDM-A (50% DMEM-F12, Thermo Fisher Scientific, #11330032; 50% Neurobasal, Thermo Fisher Scientific, #21103049; 1x B27 without vitamin A, Thermo Fisher Scientific, #12587010; 0.5x N2 supplement Thermo Fisher Scientific, #17502048; 1:4000 insulin solution, Sigma-Aldrich, I9278-5ML; 1x GlutaMAX, Thermo Fisher Scientific, #35050038; 0.5x MEM-NEAA, Thermo Fisher Scientific, #11140050; 1x Penicillin-streptomycin, Thermo Fisher Scientific, #15070063; 0.5µg Amphotericin B, Thermo Fisher Scientific, #15290018, 175µL in 500mL of 1:100 2-mercaptoethanol, Sigma-Aldrich, #8057400005). After 48 hours, media was replaced with fresh CDM-A.

3.3.3 RNA extraction, RNA-sequencing and data preprocessing

To reduce variability across replicates, we pooled organoids (Table 3.1) from half a 96 well plate from different wells, as one replicate. The organoids in each replicate were pelleted at 200G for 5 minutes, and the supernatant was removed before washing once in DPBS. RNA from two replicates was extracted for each individual time point. The cell pellet was supplemented with 350µL per sample of Buffer RLY with 3.5µL β-Me and homogenized by vortexing vigorously. RNA extraction was performed using the Bioline Isolate II RNA Mini Kit (Bioline, #BIO-52071) according to the Bench-Top Protocol by the manufacturer. Assessment of RNA integrity and purity was performed using TapeStation 4200. Samples were frozen and sent on dry ice to GeneWIZ. At GeneWIZ, total RNA of each sample was quantified and qualified by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), NanoDrop (Thermo Fisher Scientific) and 1% agrose gel. 1 µg total RNA with RIN value above seven was used for following library preparation. Next generation sequencing library preparations were constructed according to the manufacturer's protocol (NEBNext®, UltraTM RNA Library Prep Kit for Illumina®). The poly(A) mRNA isolation was performed using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB) or Ribo-ZeroTM rRNA removal Kit (Illumina). The mRNA fragmentation and priming was performed using NEBNext First Strand Synthesis Reaction Buffer and NEBNext Random Primers. First strand cDNA was synthesized using ProtoScript II Reverse Transcriptase and the second-strand cDNA was synthesized using Second Strand Synthesis Enzyme Mix. The purified double-stranded cDNA(by AxyPrep Mag PCR Clean-up (Axygen) was then treated with End

Prep Enzyme Mix to repair both ends and add a dA-tailing in one reaction, followed by a T-A ligation to add adaptors to both ends. Size selection of Adaptor-ligated DNA was then performed using AxyPrep Mag PCR Clean-up (Axygen), and fragments of 360 bp (with the approximate insert size of 300 bp) were recovered. Each sample was then amplified by PCR for 11 cycles using P5 and P7 primers, with both primers carrying sequences which can anneal with the flow cell to perform bridge PCR and P7 primer carrying a six-base index allowing for multiplexing. The PCR products were cleaned up using AxyPrep Mag PCR Clean-up (Axygen), validated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), and quantified by Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). Then libraries with different indices were multiplexed and loaded on an Illumina HiSeq instrument according to manufacturers instructions (Illumina, San Diego, CA, USA). Sequencing was carried out using a 2x150bp paired-end (PE) configuration; image analysis and base calling were conducted by the HiSeq Control Software (HCS) + OLB + GAPipeline-1.6 (Illumina) on the HiSeq instrument. The sequences were processed and analyzed by GENEWIZ. Demultiplexing was performed by bcl2fastq 2.17. Raw data was filtered as follows: Pair-end reads with adapter, pair-end reads when the content of N bases is more than 10% in either read and pair-end reads when the ratio of bases of low quality ($Q<20$) is more than 0.5 in either read, were discarded.

Paired-end 150 bp reads were mapped to the human (GRCh38) genome using STAR (v2.7.9a) [134] allowing up to three mismatches per pair, retaining only reads that are aligned to a unique location, and permitting a maximum intron length of 100,000. Alignments containing non-canonical junctions were filtered. Annotated transcripts (GRCh38 release 104) were provided to extract information on splice junctions to improve the accuracy of the alignment. Reads were summarized to quantify Ensembl/Refseq-annotated exons using Subread (v1.4.6) [135]. Batch correction for replicates was performed using ComBat_seq (v3.36.0) [136]. Reads per library are summarised in Figure 3.2E. To visualize the log2FC across time points and to perform PCA and unsupervised hierarchical clustering of the RNA-seq data, regularized log transformation of read counts were used as implemented in DESeq2 [137].

Table 3.1. Number of organoids used for RNA-seq. Total number of organoids per replicate quoted in the table.

Time Point	RNA-seq (n=2)
Day 0	96
1 hour	48
6 hours	48
Day 1	24
Day 2	24
Day 4	10
Day 6	10
Day 8	10

3.3.4 Sample and library preparation for scRNA-sequencing and data preprocessing

20 organoids were collected at day 8 of differentiation in a 1.5 mL tube and washed twice with PBS. 1 ml of complete Neurosphere Dissociation Kit (Miltenyi, #130-095-943) was added and the tube was kept in a 37 °C water bath for 5 minutes. After 5 minutes the tube was flicked and inverted several times to encourage the organoids to break up and placed at 37 °C for another 5 minutes. The suspension was then slowly pipetted up and down 10 times with a 1000µL pipette and the tube was placed for a further 5 minutes in the water bath. Following this incubation, the suspension was slowly pipetted up and down 10 times with a 200µL pipette and passed through a 30µm cell strainer (Miltenyi, #130-041-407) to remove any large aggregates. 500µL of HBSS was passed through the cell strainer and added to the cell suspension, which was then centrifuged at 200g for 5 minutes. The pellet was resuspended in 200µL PBS + 0.04% BSA prior to a live cell count. Sample single cell suspension viability and concentration was assessed using Trypan Blue staining on a Countess II Automated Cell Counter. Cell suspension was loaded on a 10x Genomics Chip G for a target output of 10,000 single cells. Single cell libraries were prepared in accordance with Chromium Single Cell 3' Reagent Kit v3.1 Dual Index user guide. Samples were sequenced on an Illumina NovaSeq according to manufacturer's instructions, using 2x150 paired-end configuration, and a depth of 200M reads/sample. Image analysis was performed in real time by the NovaSeq Control Software (NCS) v1.2.0.28691 and real-time base calling was performed using Real Time

Analysis v4.6.7. Then the Illumina DRAGEN BCL Convert 07.021.645.4.0.3 pipeline was used to generate the sequence data.

Counts were generated using 10X Genomics Cell Ranger (v7.2.0) [138] and UMIs were mapped to the GRCh8 genome. Empty droplet filtering was performed using DropletUtils (v1.22.0) [139]. After standard cell filtering, doublet detection was performed using DoubletFinder (v2.0.4) [140]. After data was normalised and scaled as per the Seurat (v5.0.1) [41] pipeline, shared nearest-neighbour graph construction and unsupervised clustering were performed using *FindNeighbours()* and *FindClusters()* respectively. Resulting clusters were manually annotated using known marker genes.

3.3.5 Organoid sample preparation for mass spectrometry

Organoids (Table 3.2) were homogenized and lysed with a drill and pestle in 160 μL 50mM HEPES pH 7.4, 2mM EGTA, 2mM EDTA, 2mM phenylmethylsulfonyl fluoride, EDTA-free Protease Inhibitor (Roche), and PhosSTOP (Roche). 40μL of 10% w/v SDS was added for a final concentration of 2% w/v SDS and the lysate was frozen on dry ice. Frozen lysate was thawed and reduced with 10mM tris(2-carboxyethyl)phosphine (TCEP) at 85 °C for 10 minutes with shaking. Samples were alkylated with 20mM iodoacetamide for 30 minutes at 23 °C in the dark. Protein was precipitated using the chloroform-methanol method and pellet reconstituted in 20μL of 7.8M urea and 50 mM HEPES (pH 8.0). Protein was digested with 5μg Lys-C (FUJIFILM Wako Pure Chemical Corporation) and incubated for 8 hours at 25 °C, with shaking. Samples were diluted 8-fold with 50mM HEPES (pH 8.0) and digested by the addition of 5μg trypsin (TrypZean, Sigma-Aldrich) and incubation for 8 hours at 30 °C, with shaking. Approximately 140μg of each of the 40 samples [MG1] were derivatised with TMTpro reagent (Thermo Fisher Scientific) using the following scheme. Replicate 1: 0 minutes , 129C; 15 minutes, 130N; 30 minutes, 130C; 1 hour, 131N; 6 hours, 131C; 1 day, 132N; 2 days, 132C; 4 days, 133N; 6 days, 133C; and 8 days, 134N. Replicate 2: 0 minutes, 131N; 15 minutes, 130C; 30 minutes, 130N; 1 hour, 129C; 6 hours, 129N; 1 day, 128C; 2 days, 128N; 4 days, 127C; 6 days, 127N; and 8 days, 126. Replicate 3: 0 minutes , 129C; 15 minutes, 130N; 30 minutes, 130C; 1 hour, 131N; 6 hours, 131C; 1 day, 132N; 2 days, 132C; 4 days, 133N; 6 days, 133C; and 8 days, 134N. Replicate 4: 0 minutes, 131N; 15 minutes, 130C; 30 minutes, 130N; 1 hour, 129C; 6 hours, 129N; 1 day, 128C; 2 days, 128N; 4 days, 127C; 6 days, 127N; and 8 days, 126. Samples were combined within each replicate and desalted using a solid phase extraction (SPE) cartridge (Sep-Pak 3 cc Vac tC18, Waters). Samples were enriched for phosphopeptides using the 'TiSH' method [141]. Both the phosphopeptide enriched and de-enriched solutions were applied to hydrophilic interaction chromatography using a Dionex Ultimate 3000 HPLC system with a 250mm long and 1mm inside diameter TSKgel Amide-80 column (Tosoh Biosciences) at 50 µL/min [141, 142]. Fractions were collected every minute using a Probot (LC Packings) and those containing peptide, as determined by absorbance at 214 nm, were concatenated so that they had approximately equal peptide content. 13 fractions containing mainly phosphopeptides, a fraction containing mainly multi-phosphorylated peptides and ten fractions containing mainly-non-phosphopeptides were dried and then reconstituted in 0.1% formic acid for LC-MS/MS.

Table 3.2. Number of organoids used for mass-spectrometry. Total number of organoids per replicate quoted in the table.

Time Point	**LC-MS/MS (n=4)**
Day 0	96
1h	144
6h	144
Day 1	72
Day 2	72
Day 4	38
Day 6	38
Day 8	38

3.3.6 Mass spectrometry analysis of organoid samples

The LC-MS/MS was performed using a Dionex UltiMate 3000 RSLC nano system and Q Exactive Plus hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific). Each fraction was loaded directly onto an in-house 300mm long 0.075mm inside diameter column packed with ReproSil Pur C18 AQ 1.9µm resin (Dr Maisch, Germany). The column was heated to 50 °C using a column oven (PRSO-V1, Sonation lab solutions, Germany) integrated with the nano flex ion source with an electrospray operating a 2.3 kV. The S lens radio frequency level was 50 and capillary temperature was 250 °C. Each mainly phosphopeptide or multi-phosphopeptide fraction was injected in 5µL and loaded onto the column in 99% reversed phase buffer A (solution of 0.1% formic acid) and 1% buffer B (solution of 0.1% formic acid, 90% acetonitrile) for 25 minutes at 300 nL/minute. The gradient, at 250 nL/minute, was from 1% buffer B to 5% buffer B in 1 minute, to 25% buffer B in 74 minutes, to 35% buffer B in 8 minutes, to 99% buffer B in 1 minute, held at 99% buffer B for 2 minutes, to 99% buffer A in 1 minute and held for 8 minutes as the flow rate ramped up to 300 nL/minute. MS acquisition was performed for the entire 120 minutes. All samples and fractions were analysed using data-dependent acquisition LC-MS/MS. The MS scans were at a resolution of 70,000 with an automatic gain control target of 1,000,000 for a maximum ion time of 100 ms from 375 to 1500 m/z. The MS/MS scans were at a resolution of 35,000 with an automatic gain control target of 200,000 and maximum ion time of 115 ms. The loop count was 12, the isolation window was 1.1 m/z, the first mass was fixed at m/z 120 and the normalized collision energy was 34. Singly charged ions and those with charge >8 were excluded from MS/MS and dynamic exclusion was for 35 s. Each mainly non-phosphopeptide fraction was injected in 3.5 µL and loaded onto the column in 99% reversed phase buffer A and 1% buffer B for 17.5 minutes at 300 nL/minute. The gradient, at 250 nL/minute, was from 1% buffer B to 6% buffer B in 1 minute, to 28% buffer B in 71.5 minutes, to 35% buffer B in 8 minutes, to 99% buffer B in 1 minute, held at 99% buffer B for 2 minutes, to 99% buffer A in 1 minute and held for 8 minutes as the flow rate was increased to 300 nL/minutes. MS acquisition was performed for the entire 110 minutes.

All samples and fractions were analysed using data-dependent acquisition LC-MS/MS. The MS settings were the same as for phosphopeptides, except that the MS/MS scans had a maximum ion time of 100 ms and the normalized collision energy was 31.

3.3.7 Database searching of MS/MS data and data preprocessing

The raw mass spectrometry data files were analysed with MaxQuant 1.6.7.0 using a *Homo sapiens* fasta file downloaded on July 19, 2021, from UniProt, containing 78,120 canonical and isoform entries. The built-in MaxQuant contaminants fasta file was included. Post-translational modification was set to true for the mono-phosphorylated and multi-phosphorylated peptide files. Variable modifications were oxidation on Met, acetylation on protein N-terminus, deamidation on Asn or Gln and phosphorylation on Ser, Thr or Tyr. Carbamidomethyl modification on Cys was a fixed modification. Digestion was specific for trypsin, but allowing cleavage before proline. Up to three missed cleavages were allowed. TMTpro correction factors for 13C were entered for lot number VB294905. TMTpro reporter tolerance was 0.005 Da. Minimum reporter production ion fraction was 0.6. Minimum peptide length was 6 amino acid residues. Maximum peptide mass was 6,000 Da. Modified peptides were excluded from use in protein quantification and unmodified counterpart peptides of modified peptides were also excluded. Second peptide search and dependent peptide search were enabled. All other settings were default, including a false discovery rate of 1% for proteins and peptides and a minimum modified peptide score of 40.

A total of 27,565 phosphosites were identified after filtering for potential contaminants and reverse matches. The phosphorylation of each site was quantified by TMT isobaric labelling and the intensity is logarithm transformed (base 2). Phosphosites that were quantified in at least 25% of all samples were retained. Additional filtering required that phosphosites were quantified in 50% of replicates in at least one time point. This resulted in a total of 15,868 phosphosites (Class I: 14,632, Class II: 1,178, Class III: 58) that adequately passed the filtering criteria. The phosphorylation of each site for each sample was adjusted by the global median and missing values were imputed using a heuristic random-tail method 80 and imputed data were batch-corrected using the PhosR package [143] before normalizing by the total proteome.

A total of 9,190 proteins were identified after filtering for potential contaminants and reverse matches. Like the phosphoproteomics data analysis, TMT isobaric labelling quantifications are logarithm transformed (base 2) and proteins that are quantified by at least 25% of all samples were retained and adjusted by the global median. Missing values were imputed using the random-tail method [143, 136] and subsequently batch-corrected using ComBat and RUV R packages [143, 136].

3.3.8 *In vitro* protein kinase assay and mass spectrometry analysis

A total of seven peptides containing five phosphorylation sites of interest and two peptides which are known substrates of AKT were synthesised (Peptide Synthesis Facility, Childrens Medical Research Institute) as 11-14mers centred on the phosphorylation sites. Some peptides contained additional C-terminal Arg residues, not present in the protein sequence, to aid y ion generation in MS/MS spectra. The sequences, gene and target phosphorylation site were as follows:

RRRVRSFISPIR	TCF20	S1259
AQRQRSGSREQAGR	JMY	S58
LQRQGSRENGSR	SALL4	S748
KSEASSSPPVVR	SOX2	S250
NHVDESPVEFPR	DNMT3B	S136

Additional known substrates of AKT were included:

RERKSSSSEDR	BRAF	S428/S429
RDRSFSEGGER	ZFP36L1	S90/S92

The seven peptides were separately incubated with AKT1/2/3 under the following conditions. The 20µL solution contained 2µM of peptide, 25mM HEPES pH 7.4, and 400µM ATP and 10mM MgCl2. 4µL of AKT1/2/3 (active, GST-tagged, human, expressed in Sf9 cells) or water was added to initiate the reaction at 30 °C for 45 min. The reaction was topped by addition of 80µL of 0.5% TFA. The samples were desalted using stop-and-go-extraction tips with additional POROS R3 material added. 20% of each sample was mixed into two groups of 5 peptides to be analysed by LC-MS/MS with minimal retention time co-elution. The retention times were determined in a preliminary data-dependent acquisition run using the same LC-MS/MS equipment as for the organoid samples. Thereafter, the peptides were detected using parallel reaction monitoring LC-MS/MS. The peptide groups were loaded in 5µL in buffer A for 25 minutes at 300 nL/min. The gradient, at 250 nL/minute, was from 1% buffer B to 6% buffer B in 0.5 minutes, to 28% buffer B in 48.5 minutes, to 35% buffer B in 4 minutes, to 99% buffer B in 1 minute, held at 99% buffer B for 2 minutes, to 99% buffer A in 1 minute and held for 8 minutes as the flow rate ramped up to 300 nL/minute. Using an inclusion list with the peptides m/z values required to detect both phosphorylated and non-phosphorylated peptides, MS/MS detection was for scheduled 5 to 11 min windows over the 90 minute run. The MS/MS resolution was 35,000, the automatic gain control target was 500,000, the maximum ion time was 250 ms, the isolation window was 1 m/z and the normalised collision energy was 28. The raw data was analysed with Skyline 20.1.0.76. A peptide was considered phosphorylated if the combined intensity, of at least four transitions within 2 ppm of the predicted product ion m/z, was higher than the intensity for the same signal in the sample where no kinase was added.

3.3.9 Treatment of hCOs

Once cells reached 60-70% confluency, cells were dissociated into single cells using warm Accutase (Thermo Fisher Scientific, #A1110501). Following centrifugation (200G, 5 min) the supernatant was discarded and cells resuspended in Essential 6 medium (Thermo Fisher Scientific, #A1516401) containing 4ng/mL FGFb (R&D Systems, #RDS233FB010CF or Peprotech, #100-18C) and 50 µM Y-27632 (Dihydrochloride) (STEMCELL Technologies, #72307) at a concentration of 60 cells/µL.

150µL/well of this suspension was seeded into an ultra-low attachment 96-well plate (Thermo Fisher Scientific, #174925). Following successful formation of hPSC aggregates after centrifugation (200G, 5min), media was changed to PIM. From this point media was changed daily with PIM, transitioning to CDM-A at days 6-8. Various concentrations of SC79 (abCam, #ab146428) were given at media changes from day 1 and various concentrations of MK2206 (Selleck, #S1078) from day 2 until collection.

3.3.10 Western blot of hCOs

Organoids were homogenized and lysed in RIPA buffer (Thermo Fisher Scientific, #89901) and protease inhibitor cocktail (Thermo Fisher Scientific, #78429). The samples were centrifuged at 17000g for 15 minutes and the protein concentration was measured using Pierce BCA Protein Assay Kits (Thermo Fisher Scientific, #23225). NuPAGE LDS sample buffer (Thermo Fisher Scientific, #NP0007) and sample reducing agent (Thermo Fisher Scientific, #NP0009) was added to 20 µg of protein and boiled at 95 °C for 5 minutes. The samples were run on NuPAGE 4 to 12%, Bis-Tris, 1.01.5 mm, Mini Protein Gel (Thermo Fisher Scientific, #NP0329) at 150 volts for 50 minutes. The gel was transferred to a methanol activated 0.45 µm PVDF membrane (Thermo Fisher Scientific, #88518) at 30 volts for 90 minutes. Transfer buffer consisted of 20% v/v methanol and NuPAGE Transfer buffer (Thermo Fisher Scientific, #NP00061). The membrane was blocked with 5% BSA (Sigma-Aldrich, #A2153), TBST (Thermo Fisher Scientific, #J60448.K3) for 2 hours and incubated overnight at 4 °C with 1:1000 anti-AKT (Cell Signalling, #9272S) or 1:1000 anti-pAKT (Cell Signalling, #9271T), and 1:1000 anti-beta-actin (Abcam, #mAbcam 8226) or 1:1000 anti-GAPDH (Abcam, #Ab9485). The blots were washed and incubated with 1:10000 IRDye®680RD Goat anti-Rabbit IgG Secondary Antibody (LI-COR, #926-68071) and IRDye®800CW Goat anti-Mouse IgG Secondary Antibody (LI-COR, #926-32210) in the dark for 1 hour before imaging using Chemidoc Touch Gel Imaging System (Bio-Rad, #1708370).

3.3.11 Immunohistochemistry

Organoids were fixed with 4% (w/v) paraformaldehyde in PBS and incubated overnight in 20% sucrose, prior to embedding in OCT. Cryosections (14 µm thick) were collected and preserved at -20 °C. Cryosections were blocked in 5% goat serum in blocking solution (1% bovine serum albumin in PBS and 0.1% Triton X) for 2 hours. Primary antibody (1:300 NeuN, 1:200 NCAD, 1:200 Oct4, 1:200 Ecad, 1:200 Pax6, 1:200 Nestin, 1:200 Sox3, 1:200 MAP2, 1:2000 Tuj-1) was incubated overnight at 4 °C. Sections were incubated with Alexa Fluor secondary antibodies (488, 546; Thermo Fisher Scientific, #A11034, #A11001, #A11003, #A11056) for 2 hours at RT, washed and counter-stained with DAPI (Sigma-Aldrich, #D9542-1MG). Secondary antibodies were used at a 1:500 dilution.

3.3.12 Image analysis

Fluorescence images of immuno-staining were captured via confocal microscopy (EVOS, Stellaris 8 Leica) at 4x objective. For fluorescent images, quantification of total area, total intensity was performed using CellProfiler (v4.2.4). Area of the organoid and total intensity values were batch normalised. p-values were calculated using a two-sided Wilcoxon rank-sum test. Representative images were contrast-corrected with Icy (v2.4.3.0). Number of organoids quantified in Table 3.3.

Table 3.3. Number of organoids used to quantify total NeuN intensity of organoids at day 8 Total number of organoids per condition and batch quoted in the table.

	Batch 1	**Batch 2**
WT (control)	3	3
MK2206	12	8
SC79	13	12

3.3.13 Kinase activity inference and kinase-substrate prioritisation

To determine kinases that were significantly up- and down-regulated at specific phases of differentiation, we grouped the time points into three stages including initial, intermediate, and late stages and applied KinasePA [144] to infer the activity of each kinase at each of these stages. The kinase substrate scoring method from PhosR [143] was used to prioritise kinase-substrate relationships in the phosphoproteomics data. Specifically, the likelihood that a kinase will regulate a given substrate is estimated by combining scores that have been calculated from the dynamic phosphorylation profiles of quantified phosphosites and the kinase recognition motif. These kinase-substrate profiles were subsequently used to construct kinase networks wherein the edges of the networks represent the correlation of the kinase-substrate profiles between each pair of kinases. A Pearson correlation > 0.95 was used to filter the low concordance kinase pairs. The activity of each kinase at each time point compared to 0 hour was assessed by testing the phosphorylation level changes of its substrates using a two-sided Wilcoxon Rank Sum test.

To identify distinctive temporal phosphorylation profiles and putative kinases associated with these profiles from the phosphoproteomics data, we next used CLUE [145], a clustering and evaluation method, to partition all dynamically regulated phosphosites into specific clusters based on their temporal profiles using fuzzy c-means clustering. For each cluster, over-representation analysis was performed to annotate kinases whose substrates are significantly enriched within that cluster using the PhosphoSitePlus database [23] as a reference. To prioritise substrates that may be phosphorylated by kinases identified in temporal clusters, the combined scores generated for kinases enriched in temporal clusters from PhosR were used to select the top 20 candidates.

3.3.14 Differential analysis of multi-omics

To determine differentially regulated phosphosites and proteins across the differentiation process, an ANOVA test was performed across all time points and a post-hoc test was performed using Limma (3.54.0) [146] for each time point. Phosphosites and proteins were deemed significantly differentially regulated if they satisfied the joint criteria of a log2FC > 0.5 and FDR-adjusted p-value < 0.05 from both ANOVA and Limma analyses. To determine differentially regulated mRNAs, the likelihood-ratio test in DESeq2 was performed and the mRNA level of a gene is differentially regulated if the joint criteria of a log2FC > 0.5 and FDR-adjusted p-value < 0.05 from the likelihood-ratio test was fulfilled.

3.3.15 Enrichment analyses of gene sets and pathways

First, to determine the enrichment of dynamically phosphorylated proteins in TFs, TCs, CRs and SEGs, annotations were collated. Annotations of known TFs, TCs and CRs were gathered from the TRRUST database [24], AnimalTFDB [147] and EpiFactors2 databases [148]. Human SEGs were downloaded from the scMerge package [149] as part of a previous publication defining SEGs [150]. To calculate the odds ratio, we calculate the odds of dynamically phosphorylated proteins belonging to either gene category against the odds of the background, which we define as the total intersecting gene species in the transcriptome and proteome. The two-sided Fishers Exact Test was performed to estimate the significance of the odds ratio, and 95% confidence intervals were reported.

For pathway analysis, genes that were concordantly up- and down-regulated on both the transcriptomic and proteomic level were identified and characterised by inferring the signalling pathways associated with their regulatory activity using the directPA R package [151]. The log2FC was taken into account when analysing the enrichment of common features within gene sets annotated to a given pathway using the Reactome database [152] as reference.

3.3.16 Gene regulatory network identification

TFs that are concordantly regulated at both the transcriptomic and proteomic level were identified using directPA [151]. Specifically, Stouffer's test was used to combine p-values of the gene expression against the protein abundance for specific temporal comparisons to assess the statistical significance of highly regulated TFs. Downstream targets of identified significantly regulated TFs were obtained from the TRRUST database [24] containing experimentally validated TF-target interactions. Taking the gene expression and protein abundance of known TFs and targets, a comparative heatmap was drawn using pheatmap [153]. To derive GRNs, TFs and targets are denoted as the nodes of the network. Edges were calculated by averaging the regularized log mRNA expression between the TF and target for the terminal time points and then min-max normalised.

3.3.17 Mapping to in vivo embryo data

The correlation analysis of regulons was conducted using pre-processed single-cell datasets obtained from a single-cell human early embryonic cell atlas [154] and Mouse Organogenesis Spatial Transcriptomic Atlas (MOSTA) [7]. Where necessary, preliminary filtering procedures to remove low quality cells were performed as per previously stated methods and counts were log-transformed using NormalizeCounts from scuttle [155]. To enhance the compatibility of our bulk RNA-seq profiles to the sparse scRNA-seq profiles, the single-cell counts were averaged across all cells of a cell type with respect to its developmental stage at which it was extracted. Next, the pairwise averaged expression was calculated between previously identified TFs and targets. Pearson correlation was calculated for each bulk time point against each cell population stratified by their corresponding developmental stage.

3.4 Results

3.4.1 Generation of a time-resolved trans-omic map of early hCO formation

To investigate the molecular programs across cell signalling and their downstream regulation during the early phases of human neural development, we adapted a previously established protocol to generate 3D hCOs from hPSCs that recapitulate the *in vivo* neural development [55] and profiled the global phosphoproteome, proteome using mass spectrometry (MS), and transcriptome using RNA-seq at matched time points during the initial phase of the neural induction process (Figure 3.1A, Table 3.1, Table 3.2). Specifically, we mapped these three omic layers across 8 days of early hCO development starting from the exit of pluripotency of hPSC aggregates to the formation of the neuroepithelium (by day 8 of differentiation), capturing the signalling events and downstream transcriptional regulatory activity that occur during the differentiation process. Immunostaining for known pluripotency factors OCT4 and ECAD detected their expression at day 0 but not at day 6 when neurogenic differentiation took place as revealed by known neural markers NCAD, MAP2, and TUJ-1 (Figure 3.1B). This reflects the transition from the exit of pluripotency towards acquisition of neural identity. Other markers including PAX6, SOX3 and Nestin further confirmed that neural differentiation has proceeded by day 8 (Figure 3.2A).

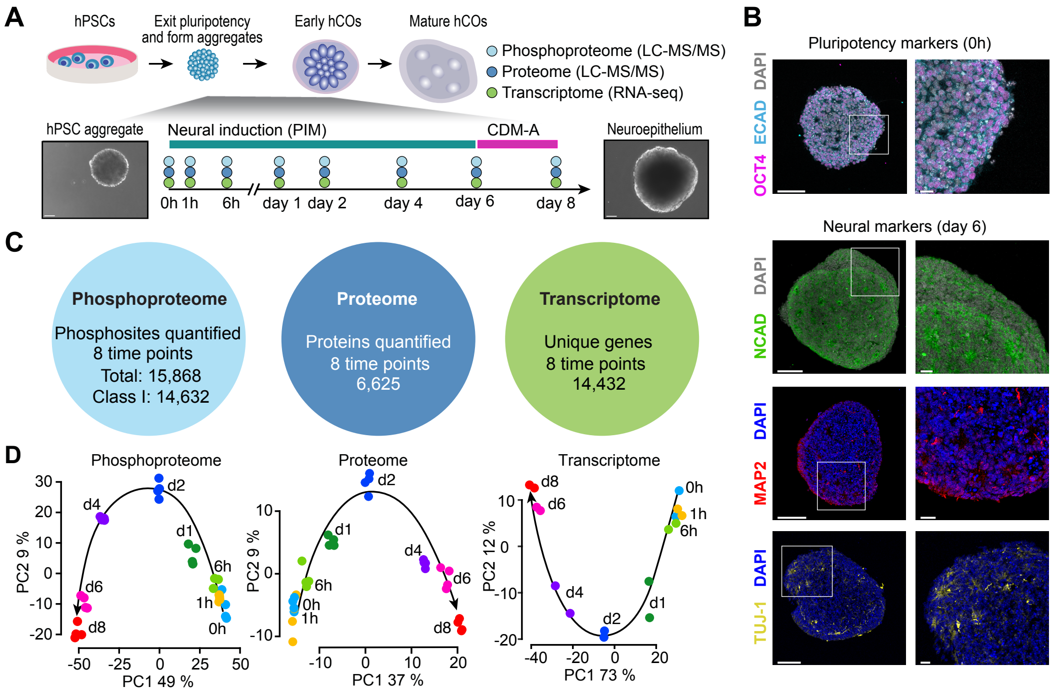


Figure 3.1. Time-resolved phosphoproteomic, proteomic, and transcriptomic profiling of early hCO formation. (A) Schematic of differentiation showing the neural induction of hCOs from hPSCs with the addition of PIM (Pro-neural Induction Media) to hPSC aggregates (day 0-5) followed by a switch to CDM-A differentiation media at day 6. The phosphoproteome (n=4), proteome (n=4) and transcriptome (n=2) were profiled at matched time points across the induction. Representative brightfield images of hPSC aggregate at 0h (left) and at day 8 (right). Scale bars represent 100µm. Organoid numbers at each time point are presented in Table 3.1 and Table 3.2. (B) Immunohistochemistry of OCT4 and ECAD at 0h and NCAD, MAP2, and TUJ-1 at day 6. Scale bars represent 100 m for whole images, and 20 m for magnified images. (C) Summary statistics of phosphosites and proteins quantified by mass spectrometry and genes quantified by RNA-seq. (D) Principal Component Analysis (PCA) showing the temporal trajectories of the phosphoproteome, proteome and transcriptome profiles during hCO formation. Biological replicates from each time point are coded by the same colour.

For the proteome and phosphoproteome, I quantified over 5,000 proteins (Figure 3.2B) and over 12,000 phosphosites (Figure 3.2C) in each of the replicates across the eight time points. This led to the quantification of 6,625 proteins and 15,868 phosphosites across the eight time points (Figure 3.2C). The majority (92.2%) of phosphosites have a localisation probability of >75% indicating a high confidence that the phosphosite is localised to a single amino acid. As expected, most phosphosites are localised on the serine residue, followed by threonine and tyrosine (Figure 3.2D). On the transcriptome level, over 15 million read counts across all replicates and time points (Figure 3.2D) were mapped to 14,432 unique genes across the time points (Figure 3.1C).

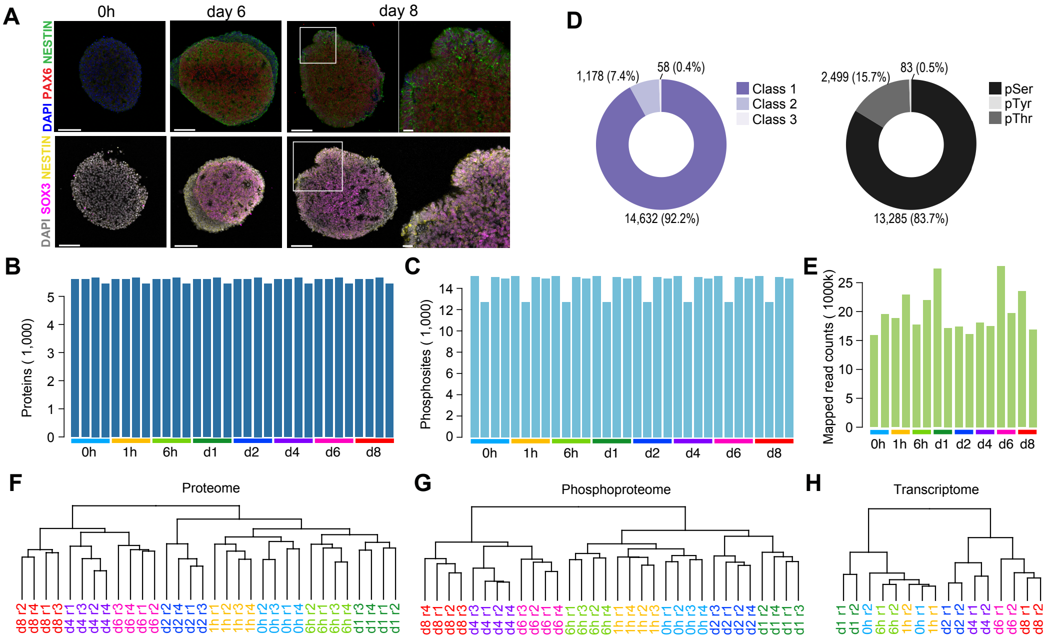


Figure 3.2. Summary statistics to show the reproducibility of replicates across the transcriptome, proteome and phosphoproteome. (A) Immunohistochemistry staining of the differentiating hPSC aggregates at 0h and organoids at day 6 and day 8. Cells were stained with DAPI, Nestin, and PAX6 (top) or SOX3 (bottom). Scale bars represent 100µm for whole images, and 20 µm for magnified images. (B, C) Numbers of proteins and phosphosites of biological replicates across time. (D) Distributions of the localisation confidence of quantified phosphosites (left). Class 1, probability >= 0.75, Class 2, probability < 0.75 and >= 0.5, Class 3, probability < 0.5 and >= 0.25. Distributions of phosphorylated amino acid residues for quantified class 1 phosphosites (right). (E) Total RNA read counts of biological replicates across time. (F-H) Unsupervised hierarchical clustering of proteomic, phosphoproteomic and transcriptomic profiles of biological replicates across time.

The PCA (Figure 3.1C) and unsupervised hierarchical clustering (Figure 3.2F-H) of the three omic layers revealed that global dynamics are dictated by the temporal molecular profiles as replicates cluster together within each time point and samples separated in a time-dependent manner. The relatively small changes in variance in 0 h, 1 h and 6 h in the PC analysis suggests a gradual transition in the first 6 h after the initiation of neural induction. However, greater changes are observed past day 1, revealing the timing of exit from the pluripotent state and commitment towards neural differentiation. Together, these data provide a trans-omic map of lineage progression from the exit of pluripotency to the neuroepithelium formation in early hCOs.

3.4.2 Kinase activity inference reveals key signalling events in controlling early hCO differentiation

To identify the key kinases and their signalling networks involved in the transition from the exit of pluripotency to the neuroepithelial formation in early hCOs, we inferred the kinase-substrate relationships from the time-course phosphoproteomic data using PhosR [143] and constructed the kinase networks based on the amount of substrate proteins they shared during this transition process. The reconstructed kinase networks reveal extensive signalling crosstalk among kinases, especially among CDKs and ERKs and those belonging to the AGC kinase family. Enrichment analysis identified a subset of these kinases (AKT, PKA, MAPKAPK2 and CDKs) are significantly regulated across the multiple time points during the differentiation process (Figure 3.3).

Figure 3.3. Inference of kinase activity network during the progression of early hCO development. Kinase networks inferred from kinase-substrate relationships derived from kinase recognition motifs and phosphoproteomic profiles during differentiation. The width of edges between kinases indicates the correlation in their kinase-substrate profiles. The outer ring shows the level of regulation of the respective kinases across time compared to 0h using a two-sided Wilcoxon Rank Sum test. * denotes those that are significantly regulated at four or more time points.

To further investigate the kinase activity and their regulation during the differentiation process, I grouped the profiled time points into initial (6 hours and day 1), intermediate (day 2 and day 4), and late (day 6 and day 8) stages and performed kinase activity inference at each stage (Figure 3.4A).

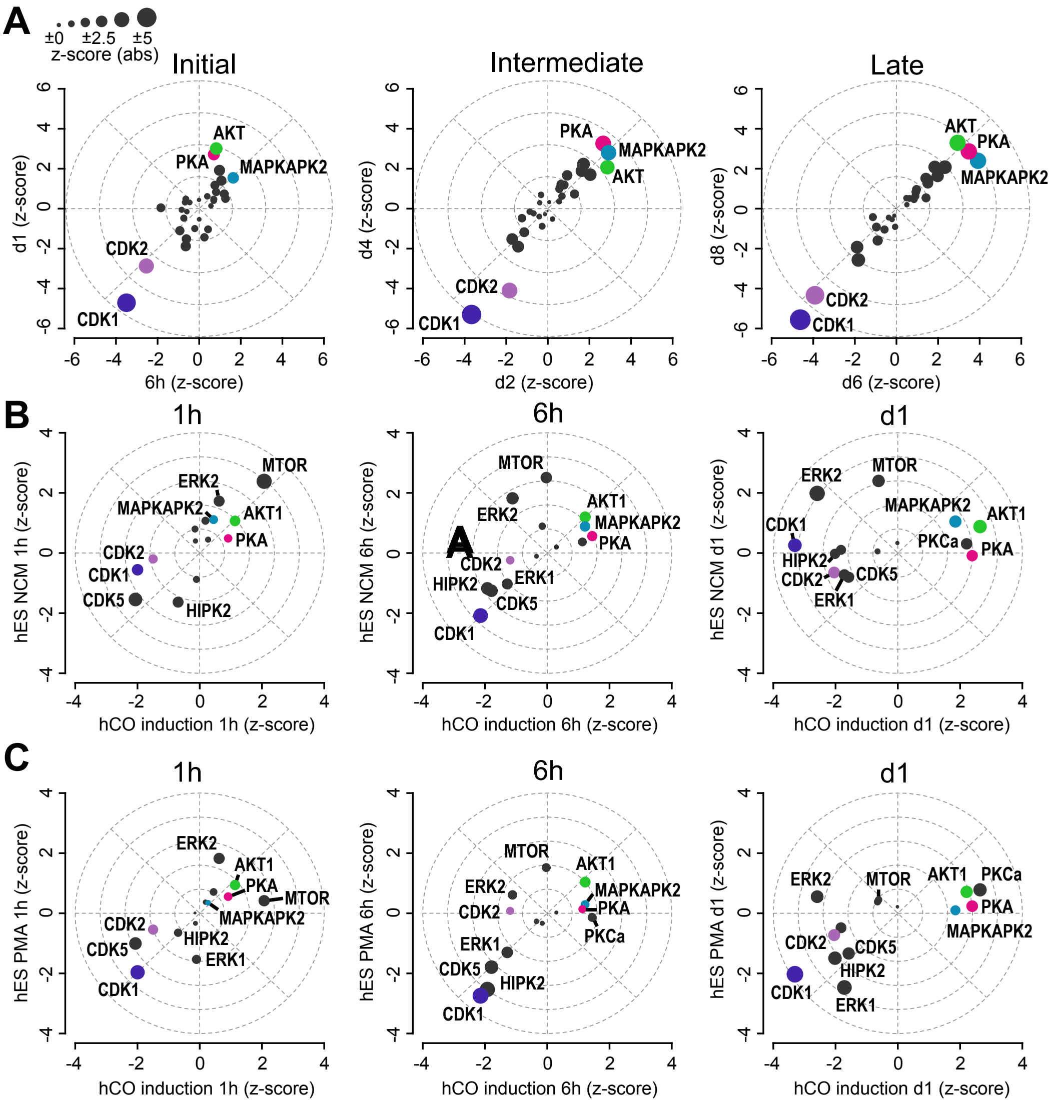


Figure 3.4. Regulation of neural-specific kinase activity during hPSC differentiation. (A) Inference of kinase activity at the initial, intermediate, and late stages during the profiled hCO differentiation process. All time points are compared to 0h. (B, C) Comparative analyses of kinase activity at 1h, 6h, and day 1 in hCO differentiation and lineage-independent hPSC differentiation induced by (B) unconditional medium (NCM) and (C) phorbol 12-myristate 13-acetate (PMA). All time points are compared to 0h.

Notably, the activity of AKT and PKA are up-regulated at the initial phase while MAPKAPK2 is up-regulated at the intermediate stage. Conversely, CDK1 and CDK2 are down-regulated continuously during the differentiation process. Comparative analyses of hCO formation phosphoproteome with those from nondirected hPSC differentiation, induced by unconditional medium (NCM) or phorbol 12-myristate 13-acetate (PMA) [156], reveal that kinase activities are most concordant at the 1 hour time point but grow increasingly discordant at later time points of 6 and 24 hours (Figure 3.4B, C), suggesting that unique combinations of kinase signalling are turned on/off in hPSC differentiation to different cell trajectories. I found that sustained AKT, PKA, and MAPKAPK2 signalling is a distinguishing feature during hCO formation. In comparison, stronger ERK2 and mTOR signalling appear to be induced by NCM and to a lesser extent PMA during lineage-independent hPSC differentiation.

Next, I sought to visualise the temporal activity of these significantly regulated kinases based on the temporal profile of their substrates. To uncover the signalling cascades driven by these significantly regulated kinases and the temporal profile of their substrates during the neuralisation process, I applied ClueR, a tool for time-course phosphoproteomic data clustering and enrichment analysis [145], and identified five distinct temporal clusters with kinase substrate enrichment (Figure 3.5).

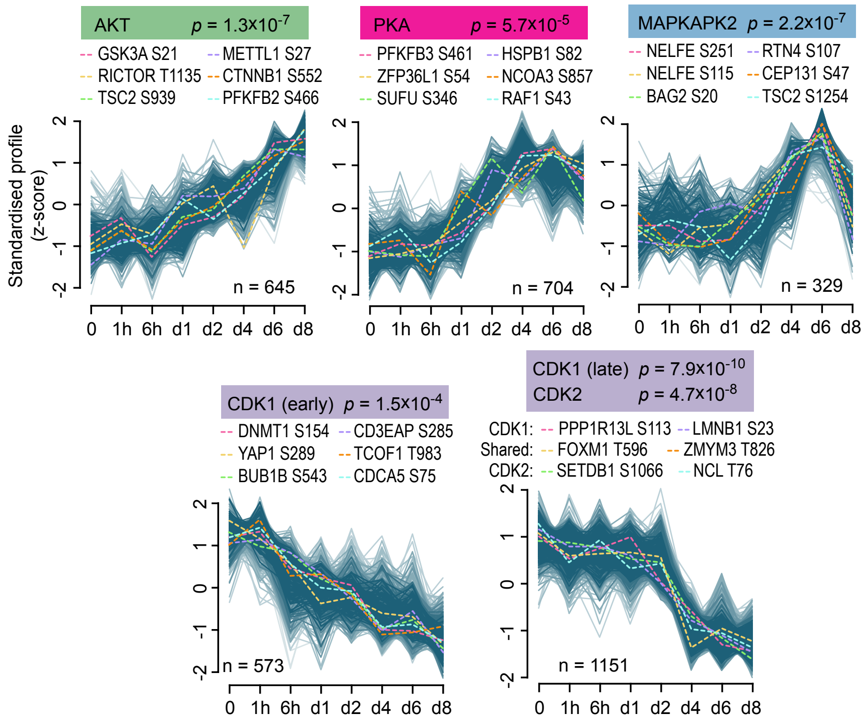


Figure 3.5. **Temporal dynamics of kinase regulation during neuralisation.** Clustering of phosphosites by their temporal dynamics. Kinases with their known substrates over-represented in each temporal cluster are shown above the cluster and representative substrates are highlighted.

In particular, I found the clusters enriched for AKT and PKA substrates are up-regulated around day 1, followed by the cluster enriched for MAPKAPK2 at day 2. In contrast, clusters enriched for CDK substrates are down-regulated during the differentiation process with some substrates down-regulated as early as 6 hours while others by day 4.

The changes in kinase activity do not necessarily couple with their changes in gene expression and protein abundance. While PKA and MAPKAPK2 show moderate up-regulation, CDK1/2 remain largely stable (Figure 3.6A). This may reflect the de-commissioning of pluripotency, as the inactivation of CDK via inhibitive phosphorylation leads to the induction of differentiation [157] independent of global changes in protein levels [158]. Notably, the three isoforms of AKT show diverse gene expression and protein abundance profiles, suggesting that they may play different roles in early hCO differentiation (Figure 3.6A). To identify substrates that may be regulated by the key kinases in the signalling cascades, I performed kinase-substrate prediction using PhosR which scores the likelihood of a phosphosite for being a substrate of a kinase based on both the temporal profile of the phosphosite and the flanking amino acid sequence surrounding the phosphosite. Figure 3.6B highlights the top list of phosphosites with a prediction score >0.7 and are associated with known transcription factors (TFs), chromatin remodellers (CRs), and transcription co-factors (TCs) in each of the five kinase clusters.

Given the specificity of AKT signalling in hCO formation compared to lineage-independent differentiation of hPSCs by NCM and PMA, I selected a panel of predicted AKT substrates for evaluation using in vitro kinase assays. These include TCF20, JMY, SALL4, SOX2, and DNMT3B, that are known to link with neural-related pathways and neural diseases (Figure 3.6B, Figure 3.7A, B). Known AKT substrates BRAFs [159, 160] and ZFP36L1s (also known as BRF1) [161] were included as positive controls. Among the putative candidates, I found that TCF20 is phosphorylated by all AKT isoforms at S1259 and JMY is phosphorylated by AKT1/3 in the in vitro kinase assays. Previous studies have reported the deletion of TCF20, which targets the downstream DNA demethylation factor TDG, is associated with autism spectrum disorder [162]. JMY, which can phosphorylated by AKT, is known to participate in the p53 response controlling the switch from cell proliferation to cell differentiation and neurite outgrowth [163]. The validation of AKT phosphorylation on these proteins points to their potential impact on early hCO differentiation at a phosphosite-specific resolution.

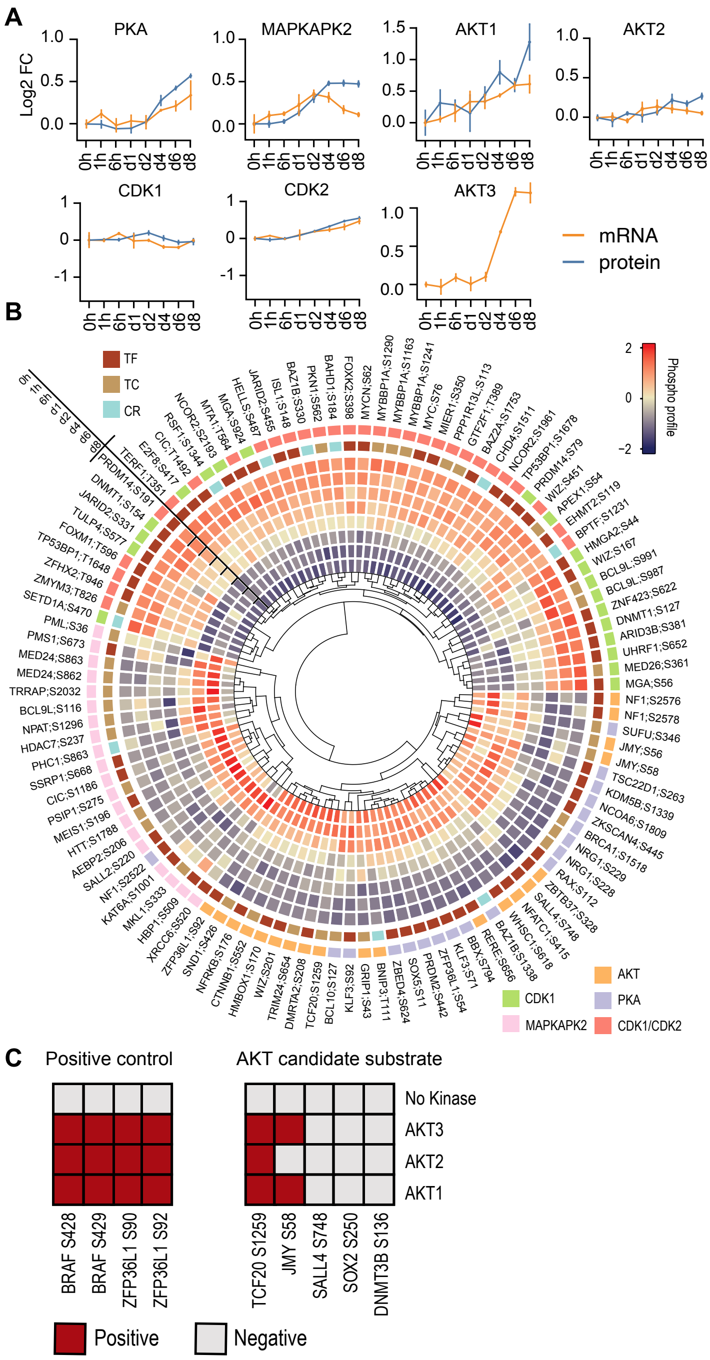


Figure 3.6. Prioritisation of kinase substrates of AKT1/2/3. (A) RNA expression and protein abundance (where available) of major kinases identified from the hCO differentiation process. Data are represented as mean ±SD. mRNA (n=2), protein (n=4). (B) Circular heatmap showing the phosphorylation profiles of phosphosites localised on predicted and known kinase substrates from PhosR. Substrates are annotated based on their respective upstream kinase regulator and whether they are classed as transcription factors (TFs), transcription co-factors (TCs) or chromatin remodelers (CRs). (C) In vitro kinase assay with human GST-tagged AKT1, AKT2, AKT3 incubated with known AKT and candidate substrates. Positive hits are indicated in dark red and negative hits are indicated in grey.

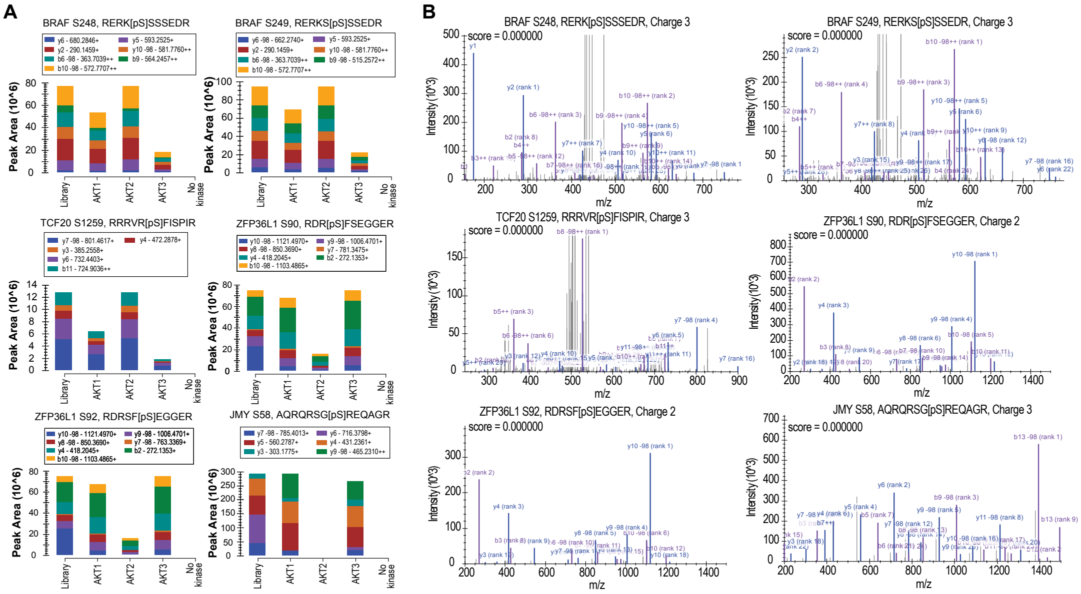
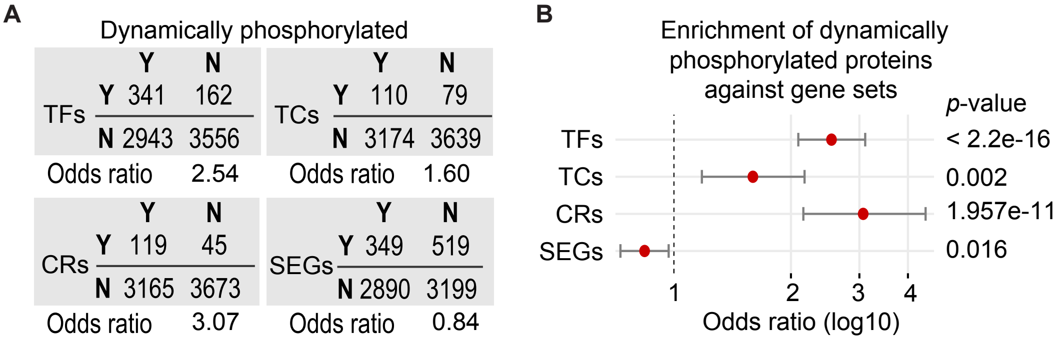


Figure 3.7. Detection of AKT1/2/3 phosphorylated peptides by parallel reaction monitoring. (A, B) The intensity values and fragment ion information obtained from the MS/MS spectra and Skyline software for each detected phosphopeptide, phosphorylated by AKT1, AKT2 and/or AKT3. (A) Shows The peak areas summed from the fragment ions, and (B) shows the MS/MS spectra and the ions that match the theoretically determined fragment ions and includes multiple positions for the phosphorylation site. All identifications were confirmed by manual inspection. Underlined amino acid residues do not exist in the protein sequence. These underlined residues, at the C-terminus, were added to improve production of y ions.

3.4.3 Phospho-signalling converges on master regulators of downstream gene regulatory networks to regulate hCO differentiation

Next, I investigated whether certain classes of gene regulatory factors were preferentially regulated by phospho-signalling during the hCO formation process. I calculated the odds ratios of dynamically phosphorylated proteins during the differentiation that are TFs, transcription co-factors (TCs), and chromatin remodellers (CRs) (Figure 3.8A). A set of stably expression genes (SEGs) [150] was included as a control for the enrichment analysis. Significance analysis using a two-sided Fishers Exact Test revealed that proteins that act as TFs, TCs, and CRs (odds ratios 2.54, 1.6, 3.07 respectively) are more likely to be dynamically phosphorylated (Figure 3.8A). In comparison, SEGs are depleted in the corresponding genes of these dynamically phosphorylated proteins. While the lack of enrichment with respect to dynamic phosphorylation of SEGs is expected as these genes are mostly involved in the housekeeping programs in cells, the enrichment of differentially phosphorylated sites, especially, on TFs and CRs reveals that these master regulators may serve as integration nodes to convert phosphorylation signals to downstream transcriptional regulation.



TFs	Y	N
Y	341	162
N	2943	3556
Odds ratio	2.54	

TCs	Y	N
Y	110	79
N	3174	3639
Odds ratio	1.60	

CRs	Y	N
Y	119	45
N	3165	3673
Odds ratio	3.07	

SEGs	Y	N
Y	349	519
N	2890	3199
Odds ratio	0.84	

Figure 3.8. Preferential phospho-signalling of regulatory factors during hCO formation. (A) Calculated odds ratios of differentially phosphorylated regulatory factors against the background (all identified proteins) compared to different gene categories. (B) Log10-transformed odds ratios and p-values calculated from a two-sided Fishers Exact Test. Error bars represent the 95% confidence interval.

To elucidate the downstream transcriptional and translational programs that drive the transition towards the neuroepithelium, I correlated the transcriptome and proteome data to identify time points which show the highest dynamics (Figure 3.9A). The pairwise correlations reveal that the correlation improves significantly at day 1, day 2 and day 8 relative to their prior time points during the transitions. To pinpoint the key factors that are involved in driving the coordinated change at these key transition points, we identified the transcriptional regulators that are significantly up- and down-regulated on both expression and protein levels using the directPA package [151] at these three time points relative to their prior time points (Figure 3.9B, C) and reconstruct the GRNs of hCO formation using the TRRUST database [24] that annotate experimentally validated TF-target interactions (3.9D).

Early neural differentiation involves the induction of pro-neural and pro-differentiation factors complemented by the loss of inhibitory factors that sustain pluripotency at as early as day 1 of differentiation (Figure 3.8D, E). This is accompanied by a significant reduction of the expression of pluripotency factors (POU5F1 and FOXO1) [164, 165] on exit of pluripotency (Figure 3.9B, C). Furthermore, the reduction in DNMT3B expression, an indicator of active priming of neural progenitors [166, 167] is accompanied by the up-regulation of neural TFs (PAX6 and SOX3) [168, 169, 170], consistent with the immunohistochemistry staining data

(Figure 3.2A) and the expression of GLI3, ZBTB16, and YAP1, which are associated with the activation of genes for diverse neural lineage specification [171, 172, 173, 174].

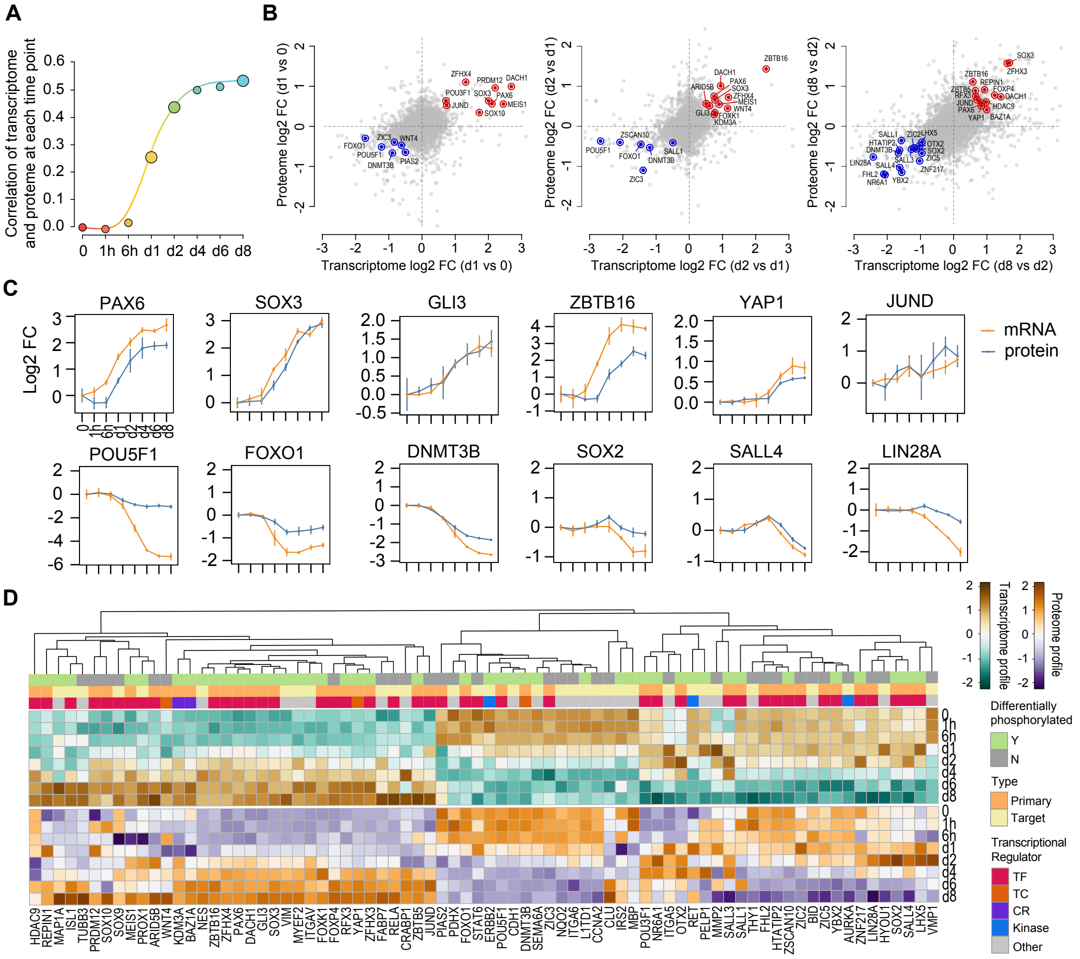


Figure 3.9. Reconstruction of gene regulatory network and characterisation of their crosstalk with phospho-signalling during hCO formation. (A) Scatter plot of time points identified as dynamically regulated on both transcriptomic and proteomic levels. (B) Scatter plot of transcriptome (x-axis) and proteome (y-axis) changes of genes in day 1 vs 0 (left), day 2 vs day 1 (middle) and day 8 vs day 2 (right). Significantly regulated transcription regulators on both the transcriptomic and proteomic levels are highlighted in red (up-regulation) or blue (down-regulation) colours. (C) mRNA and protein abundance of selected regulatory factors identified to be concordantly upregulated and downregulated across the three dynamic phases of neural differentiation. Data are represented as mean ±SD. mRNA (n=2), protein (n=4). (D) Heatmap of reconstructed GRNs showing gene expression and matched protein abundance of significantly regulated transcription regulators (denoted as "primary") and their targets as annotated in the TF-target database.

To further dissect the specific neural lineages that the above TFs and kinases contribute towards in hCOs, we performed scRNA-seq to assess the activity of up-regulated TFs and significantly regulated kinases in day 8 hCOs. Unsupervised clustering and marker annotation revealed that cells in day 8 hCOs may have diversified and correspond to neural progenitors, neural crest precursors, and multiple neuronal lineages (Figure 3.10A, B). The loss of pluripotent markers is confirmed by day 8 (Figure 3.10C) and agrees with our immunohistochemistry staining in Figure 3.1B. Notably, many of the above identified TFs show expression distinct to cell subpopulations (Figure 3.10D). Among these, PAX6 and SOX3 are strongly localised in the neural progenitor population, which is consistent with their known roles in the determination of the neuroectoderm [168, 170]. GLI3 is expressed in both the neural progenitor and neural crest precursor populations and a subpopulation of neuronal lineage. ZBTB16 and YAP1 is primarily localised in both the neural crest precursor and neural progenitor populations. As expected, JUND is broadly expressed across all cells as it serves to control cell proliferation, differentiation and survival as part of the AP-1 complex [175].

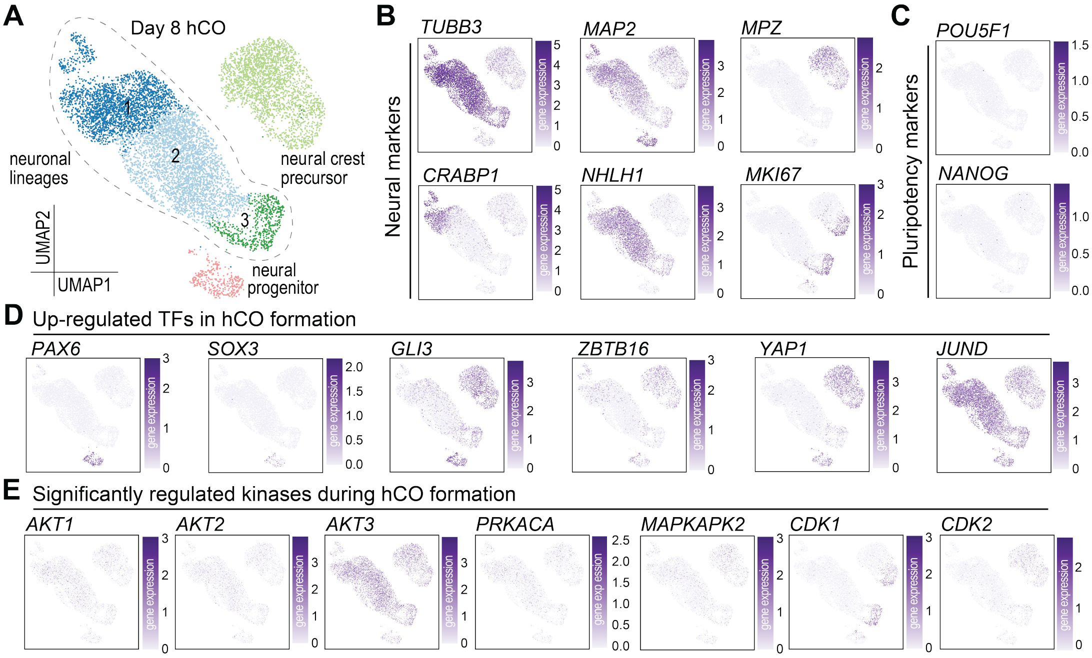


Figure 3.10. Dissect and link TFs to cell subpopulations of progenitors and neural lineages at day 8 hCOs by scRNA-seq, and identification of regulatory pathways involved during hCO formation. (A-C) UMAP visualisation of day 8 hCOs profiled by scRNA-seq (n=1). Cells are coloured by clusters determined by Seurat and annotated to subpopulations/lineages by (B) neural markers. (C) Pluripotency markers are included as controls. (D) Single-cell expression profiles of TFs that are up-regulated during early hCO differentiation (Figure 3.9C) from bulk transcriptomic and proteomic data analysis. (E) Single-cell expression profiles of significantly regulated kinases during early hCO differentiation (Figure 3.5).

In contrast to the TF profiles, most significantly regulated kinases are broadly expressed across all identified cell types (Figure 3.10E). Consistent with the bulk RNA-seq data and with literature [30], AKT3 is the most abundantly expressed isoform in the brain organoid. The lack of specificity by AKT, PRKACA and MAPKAPK2 may be expected at this early stage of neurogenesis as they are ubiquitous kinases with multiple roles regulating cell survival, metabolism, and cellular stress during the switch between neural progenitor proliferation and differentiation [32, 176, 177], Although, Levenga et al. [178] have previously reported the isoform-specific expression of AKT in specific neuronal subtypes in mice. This may lead to the possibility of kinase-specific expression developing later in human neurogenesis. Out of the significantly regulated kinases, CDK1/2 is highly expressed in a subpopulation of the neuronal lineage and a subset of neural crest precursors. This also coincides with the expression of MKI67, a typical cell proliferation marker [179], providing additional evidence that there are subpopulations of neural precursors that are actively proliferating and may be contributing to the expansion of progenitor pools prior to later specialisation. Overall, our data resource confirms existing knowledge and will be useful for discovering additional molecular factors important for neural differentiation.

Lastly, pathway analysis was performed to attribute significantly up- and down-regulated genes to pathways at each of the three stages (Figure 3.11). At day 1 compared to 0 h, there is an up-regulation of pathways related to cell replication and down-regulation of metabolism pathways associated with cellular differentiation [180]. At day 2 compared to day 1, there is a noticeable up-regulation of signalling pathways related to the PI3K-AKT-mTOR axis which is again up-regulated at day 8. This contrasts with day 1 where there is an inferred down-regulation of CDK cell cycle-related pathways which likely corresponds to the downplay of cellular replication in favour of pathways promoting differentiation. These downstream analyses of gene pathways closely align with the kinase signalling analysis and reaffirms that upstream phosphorylation plays a key role in transducing signals to transcriptional and translational regulation to induce the neurogenic program.

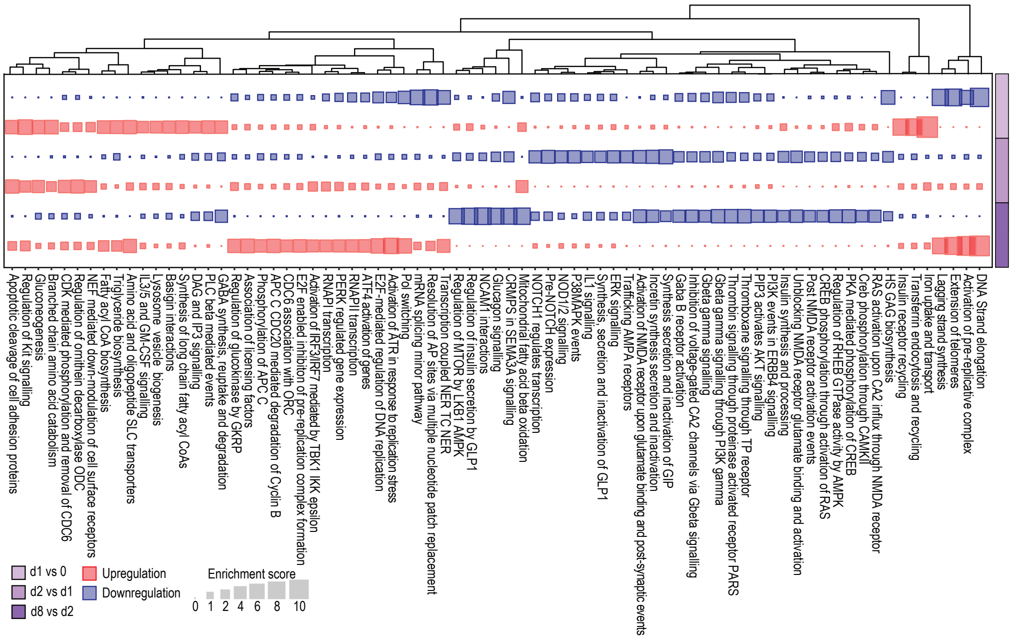


Figure 3.11. Identification of regulatory pathways involved during hCO formation. Pathway analysis using the Reactome database identifies directional regulation of signalling events by the transcriptome and proteome.

3.4.4 Comparative analysis of early hCO with developing embryos and biochemical modulation of AKT signalling for regulating early hCO differentiation

To assess whether the early hCOs mimic the in vivo counterpart of developing brains and to determine which stages of brain development they represent, I took the GRNs reconstructed from hCO differentiation process (Figure 3.12) and compared their gene expression profiles with those during human and mouse embryogenesis. Specifically, for human samples, I correlated the pairwise gene expression profiles of the GRNs during hCO formation at each time point with Carnegie stages (CS) 12-16 foetuses from a single-cell human early embryonic cell atlas [154] (Figure 3.13A).

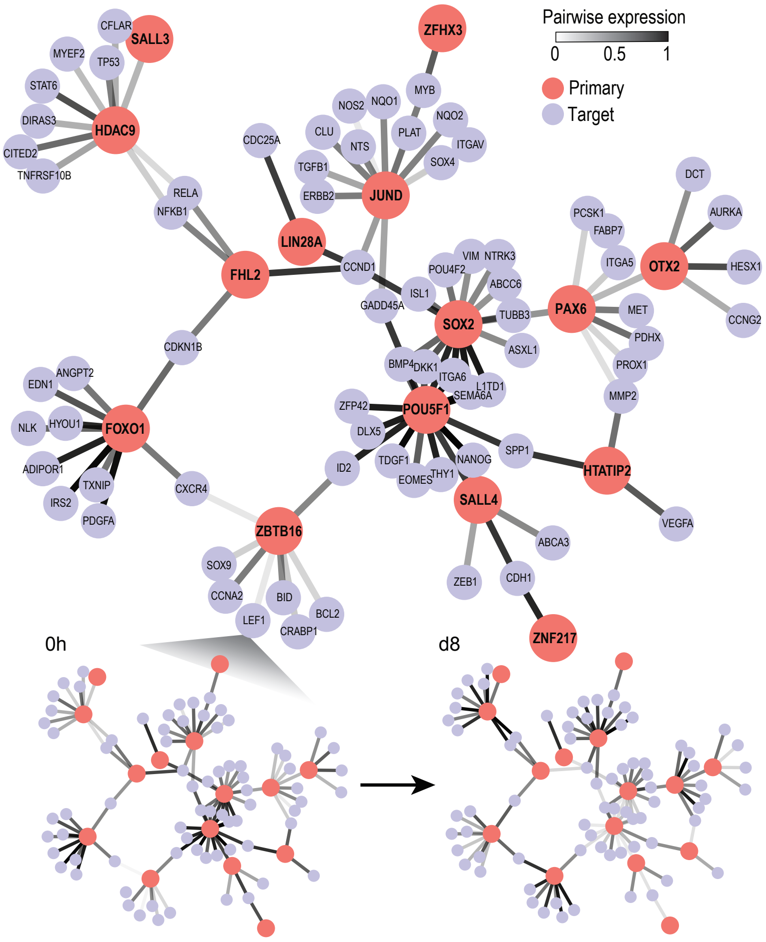


Figure 3.12. **GRN reconstructed from *in vitro* hCO differentiation process.** The transcriptional dynamics of the GRNs at 0h and d8 and are shown with edges representing the transformed averaged expression between nodes at these time points. Primary TFs are coloured in red and their target genes are coloured in purple.

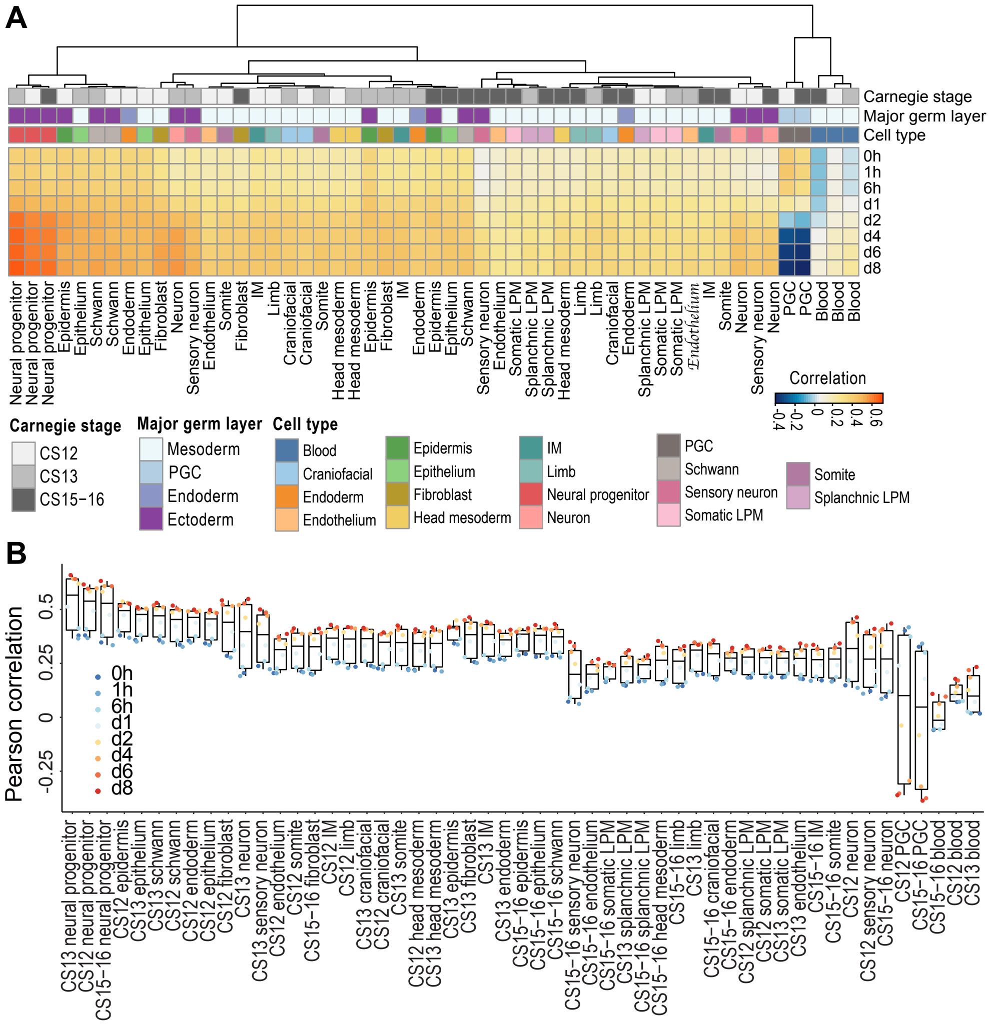


Figure 3.13. Early hCO fidelity evaluation through comparison of transcriptomic profiles with developing human embryogenesis. (A) Correlation heatmap of the gene expression profiles from early hCOs at each time point (x-axis) and their averaged scRNA-seq expression profiles in each cell type identified across human embryos at CS 12-16 (y-axis). (B) Boxplot representation of the heatmap in (A) showing correlation of the cell types at each respective human Carnegie stage compared against each time point corresponding to in vitro hCO formation. Boxplots are coloured by median correlation value; center line, median; box limits, upper and lower quartiles; whiskers, 1.5 times the interquartile range; points, outliers.

From day 2 onwards, I found an increasing high correlation of hCOs with neural progenitors in CS 12-16 (Figure 3.13A) and the overall correlation is highest with CS13 and peaks by day 8 of hCO differentiation (Figure 3.13B). Relative to other lesser correlated cell types, this suggests that the cells acquire a neural-like identity in the day 8 hCOs more alike to neural progenitors at the developing brain at CS13. The detection of an intermediate correlation with the primordial germ cell (PGC) population during initial differentiation is consistent with their association with the epiblast structure. However, as neural development progresses, there is a sharp contrasting negative correlation with the PGC population coinciding with the increasing positive correlation with the neural progenitor. As PGCs later develop into germline cells which are vastly different to the neural cell types, this may explain the negative correlation.

Next, I compared the GRNs of our early hCO differentiation data to the mouse embryo spatial atlas [154]. For each spatial location denoted by a spot in the image, I correlated its expression profile with those during hCO formation (Figure 3.14A, B).

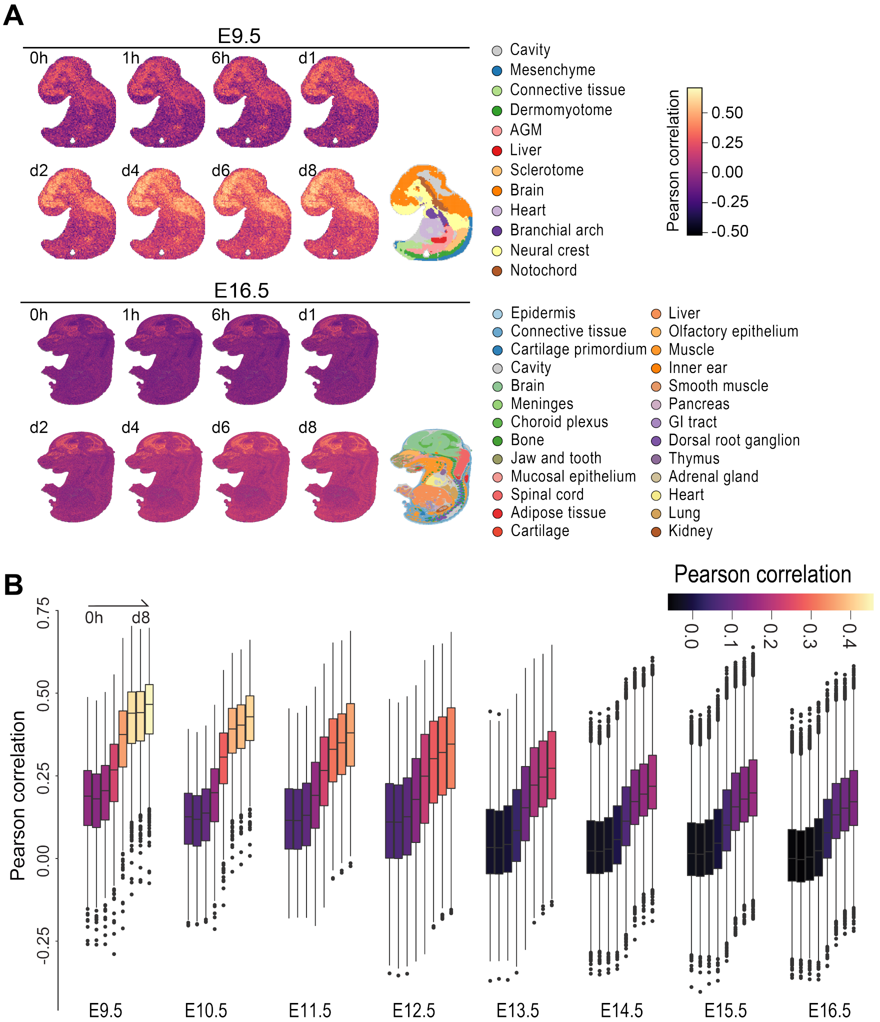


Figure 3.14. Early hCO fidelity evaluation through comparison of transcriptomic profiles with developing mouse embryogenesis. (A) Correlation of the gene expression profiles from early hCOs at each time point and their spatial transcriptomic expression profiles in the mouse embryos at E9.5 and E16.5. Correlation values were superimposed upon the spatial coordinates. Tissue regions are annotated for comparison. (B) Correlation of each hCO formation time point with the cells annotated to be in the developing brain regions during mouse embryogenesis. Boxplots are coloured by median correlation value; center line, median; box limits, upper and lower quartiles; whiskers, 1.5 times the interquartile range; points, outliers.

I found that, compared to other embryonic stages, data from the E9.5 embryo correlate the most with those from hCO differentiation and peak with the hCOs at day 8 (Figure 3.14A). In E9.5, growing correlation in regions mapping to the brain and the notochord is observed. Although the highest correlation with the brain region corresponds to E9.5 (Figure 3.14A, B), the broad correlation with other developing neural cell populations is also expected. While there is also noticeable correlation localising to the brain region at E16.5, this is to a lesser magnitude in comparison to E9.5, suggesting that cell populations in hCOs are more closely related to those at the earlier developmental stage of E9.5. Taken together, these analyses suggest that relative to other cell and tissue types in the developing embryos the in vitro early hCOs closely resemble those from embryonic brain development.

The analysis of phosphoproteomics data revealed that active AKT signalling is a unique feature in early hCO formation compared to lineage-independent hPSC differentiation and may be key for the growth of hCOs. I exploited this mechanism to improve the efficiency of the current hCO differentiation protocol by biochemical modulation of AKT activity using SC79, a small molecule AKT enhancer [181]. MK2206, an allosteric inhibitor of AKT was included as a negative control. We performed a western blot on organoids treated with SC79 and MK2206, confirming that SC79 treatment increased phosphorylated AKT levels, indicating greater AKT activation, while MK2206 led to a complete loss of phosphorylated AKT (Figure 3.15A).

By tracking the organoid growth at day 8, we found that the expansion in size of MK2206 treated organoids was significantly restricted (Figure 3.15B, C). In contrast, SC79 treated organoids were bigger in size than those from WT treatment by day 8. To further investigate whether a growth phenotype was accompanied by enhanced maturation of hCOs, I also quantified the normalised total intensity of NeuN, a neuronal differentiation marker [182], at day 8 (Figure 3.15D). As expected, inhibition of AKT activity by MK2206 led to significantly lower quantification of NeuN. In contrast, the addition of SC79 in the differentiation protocol led to a significantly higher NeuN intensity compared to WT organoids. Stainings of additional neurogenic markers including PAX6 [183] and MAP2 [182] further demonstrate these findings (Figure 3.15E). Together, these results demonstrate that modulating AKT signalling by a small molecule has a significant effect on hCO formation and neurogenic differentiation.

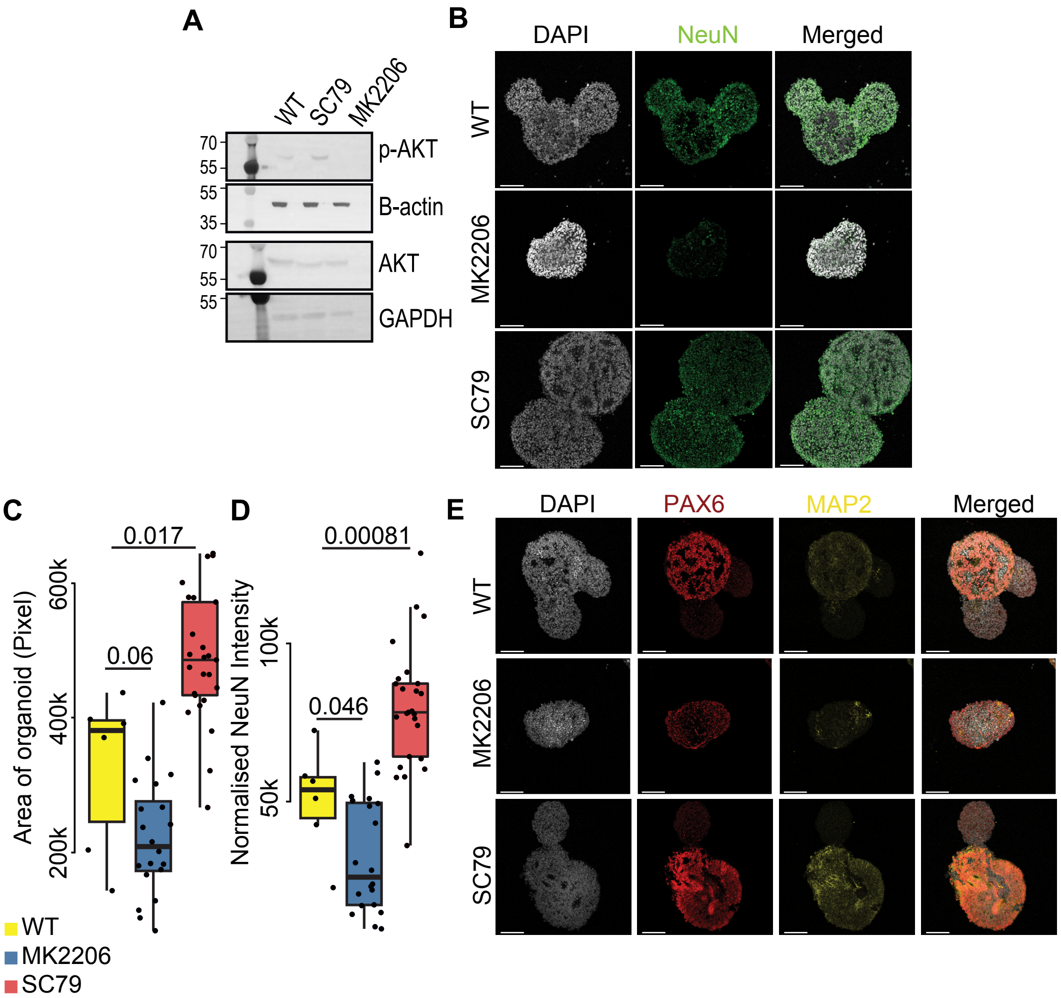


Figure 3.15. Efficacy evaluation of SC79 to accelerate neuralisation for hCO differentiation. (A) Representative western blot displaying expression levels of AKT and p-AKT in organoids at day 8 treated with SC79 and MK2206, compared with the WT. (B) DAPI and NeuN immunohistochemistry staining of the organoids at day 8 after treatment with MK2206 and SC79, compared with WT. Scale bars represent 100µm. WT, budding organoid; MK2206, single organoid; SC79, fused organoid. (C) Boxplots of the area of the organoids and (D) Normalised total NeuN intensity at day 8 in the field of view (FOV) after treatment with MK2206 and SC79, compared against WT. p-values are reported using a two-sided Wilcoxon rank-sum test. Organoid numbers for each batch are reported in Table3.3 For all boxplots: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5 times the interquartile range; points, outliers. (E) DAPI, PAX6, and MAP2 immunohistochemistry staining of the organoids at day 8 after treatment with MK2206 and SC79, compared with WT. Scale bars represent 100µm. WT, budding organoid; MK2206, single organoid; SC79, budding organoid.

3.5 Discussion

Previous characterisation studies of hCOs have been mainly limited to a single omic layer, measuring the transcriptomes, and to a lesser extent, the proteome landscape. Our study provides time-resolved phosphoproteomics data, mapping the phospho-signalling landscape that underlie the exit of pluripotency to early hCO formation. The generation of time-matched transcriptomics and proteomics data enables integrative analyses that uncover a hierarchy of intermolecular interactions occurring between cell signalling and transcriptional regulations. This broadens the molecular hierarchy among GRNs [34, 184], and highlight the quintessential role of kinases and their phosphorylation on downstream regulators such as TFs, CRs and other co-factors to drive neurogenesis.

While significant changes in phosphorylation levels were observed in many TFs and CRs during neurogenesis (Figure 3.8A), only a small subset of them show significant changes in their gene expression and protein abundance during this process (Figure 3.8B). This finding suggests that phospho-signalling may not necessarily translate into a change in gene expression and/or protein abundance. Alternatively, it may affect other molecular programs such as protein-protein interactions and sub-cellular localisation [12]. It can also affect the (co)-localisation of TFs and CRs, and their access to genome-wide binding locations as well as affinity to chromatin [185], which subsequently induces widespread transcriptional regulations on downstream target genes.

The analysis identified AKT as a key kinase in regulating neural differentiation and early hCO formation. This is in line with previous findings that the disruption of AKT signalling is associated with neurodevelopmental disorders and neurodegenerative diseases [31, 32]. Interestingly, the data also revealed that the three isoforms of AKT show differential gene expression and protein abundance with AKT2 remaining stable and AKT1 gradually increasing across the differentiation process, and AKT3 showing asharp increase from day 2 onwards (Figure 3.6A). Indeed, AKT3 has been previously reported to be highly localised in all areas of the human fetal brain and mouse adult brain [30], particularly as the most influential isoform in the hippocampus and cortex [186] primarily responsible for neuronal growth. Knock-outs of AKT3 have been found to affect neural development, causing microcephaly [30].

Computational prediction of AKT substrates further uncovered those that may contribute to the progression of neural differentiation. This led to the identification of TCF20 and

JMY which were subsequently validated to be positive substrates of AKT signalling (Figure 3.6B, C). Given the biological implications of these downstream substrates in regulating neural development, the identification and validation of these phosphorylation sites by AKT signalling serve as grounds to further investigate the interplay between AKT isoforms and these substrates in hCO differentiation, and by extension, early human neurogenesis. Previous studies have reported the development and use of AKT isoform-specific antibodies [30], which could be used to confirm the isoform specificity of JMY by co-immunoprecipitation of AKT isoforms, followed by probing for JMY in hCOs via a western blot over a timecourse.

While this work centered on predicting and validating transcription factors, co-factors, and chromatin remodelers, it is important to recognize that neurotrophic factors and signaling molecules, such as Wnt, Notch, and BMP [187], also play pivotal roles in regulating neurogenesis. Leveraging the data generated here, future research could involve predicting the phosphorylation of substrates which are key molecular players that drive the signaling cascades initiated by BDNF or NGF and other signalling pathways. In addition to *in vitro* kinase assays, which provide proof-of-principle that a specific kinase can phosphorylate the substrate peptide, further experiments can be conducted using a CRISPR-engineered cell line where the substrate gene is knocked out. These cells can be used to grow cerebral organoids over a timecourse, and site-directed mutagenesis can be performed to mutate the active site of phosphorylation followed by rescue experiments with phospho-mimetic mutations to explore the causal impact of kinase-substrate interactions on neurogenesis. This approach could potentially uncover novel therapeutic targets for modulating neurogenesis by demonstrating a phenotypical change as a direct outcome of altering kinase-substrate interactions.

The fidelity of the hCOs to their *in vivo* counterparts is critical for their use in brain modelling. Through comparative analyses of gene expression profiles with published transcriptomics atlases of human and mouse embryogenesis, we showed that the cells in the early hCOs are most correlated with the neural progenitors in the developing brains (Figure 3.13, Figure 3.14). These results indicate the fidelity of the early hCOs and the quality and utility of the trans-omic map derived from them as a resource for future investigation of early brain development.

Targeting AKT signalling is a logical strategy for regulating early hCO formation given the role of AKT in cell growth, proliferation, and survival across many cellular systems, including brain development. A previously study showed that the activation of AKT by genetic deletion of PTEN and lentivirus transduction of constitutively active AKTs in hESCs can enhance their differentiation to hCOs with greater size and surface complexity by week 6 of culture [188]. To bypass the need of labour-intense genetic manipulation of the AKT pathway, we took the approach of biochemical-based modulation of AKT activity to enhance organoid growth and neurogenesis (Figure 3.15). These results provide an orthogonal approach for modulating AKT signalling activity for enhancing hCO differentiation. Their further evaluation, including cross-talk with other kinase signalling pathways such as those identified in our trans-omic data (e.g. PKA, MAPKAPK2), will be useful for future development and optimisation of hCO growth and differentiation.

3.6 Summary

In Chapter 3, I integratively characterised the phospho-proteome, proteome, (sc)-transcriptome of neural organoid development from the exit of pluripotency to the stage of neural lineage commitment as a model for early neural development. Due to the ethical constraints of acquiring human embryo tissue at such an early stage, this necessitates the use of organoid models to study early neurodevelopment throughout a timecourse to understand the temporal regulation of phospho-signalling and downstream transcriptional changes. I assessed the transcriptional fidelity of the organoid model to *in vivo* human and orthogonally, mouse embryo development, where I additionally utilised spatial transcriptomics to show the neural regulon activity is highly correlated to the mouse brain region. I took advantage of this data resource to identify molecular factors, in our case, AKT, that could be further enhanced to increase the maturation efficiency of current cerebral organoid protocols. I also demonstrated the capability of the data to predict novel kinase substrates which could be important for neural development. While we performed an *in vitro* kinase assay which showed positive phosphorylation for some substrates, this does not take into account the native protein structure which is an inherent limitation of this study and further experimentation can be performed to demonstrate the kinase substrate interactions in the native organoid context.

CHAPTER 4

Spatial analysis of retinal cell fates using stem cell-derived retinal organoids

---

This work is part of a manuscript, tentatively titled:

Lim, BY., **Chen, C.**, Fredericks, A., Nilli, E., Mora, SM., Weatherstone, M., Aryamanesh, N., Newman, P., Zreiqat, H., Tam, PPL., Yang, P., Gonzalez-Cordero, A. (2024) Elucidating the Role of Retinoic Acid Signalling in Determining Human Retinal Cell Fate. *In preparation*.

My contribution: I pre-processed and analysed all the spatial transcriptomics and scRNA-seq data presented in this study. I also generated all the visualisations and interpreted all the data presented.

4.1 Motivation and overview

Building on our work of neurogenesis characterisation in Chapter 3, we shift our focus in this chapter to the organoid model to study the multi-faceted nature of human retinal development. In particular, I explore the application of spatial transcriptomics to uncover the spatial organisation during retinal organoid development and how retinoic acid (RA) signalling influences retinal cell fate. The RA signaling pathway is known to play a crucial role in eye morphogenesis and retinogenesis, however, the system-wide effects on regional delineation of the eye, retinal maturation, and cell type determination is not well understood. While many studies have used single-cell transcriptomics to examine normal retinal organoid development, and some have explored epigenetic landscapes with spatial transcriptomics or low-throughput in situ FISH, the work presented in this thesis combined multi-resolution spatial transcriptomics with single-cell RNA-seq data at key stages of retinal development under RA perturbation. This approach allows us to characterize changes in maturation signatures, cellular heterogeneity, and how retinogenesis is reflected in the spatial organization of the organoid under altered RA signalling activity.

4.2 Introduction

Retinogenesis is the process by which retinal progenitor cells (RPCs) differentiate in a temporally regulated sequence into the seven major retinal cell types that form the retinal laminae, responsible for light detection and vision [66, 189]. This process is broadly divided into two phases: an early phase, during which RPCs differentiate into ganglion cells, horizontal cells, and cone photoreceptors, followed by a late phase, where RPCs give rise to amacrine cells, rod photoreceptors, bipolar cells, and Müller glia.

RA signaling is essential for proper eye morphogenesis and dorsoventral patterning, which is crucial for anterior eye formation in the neural plate [190, 64]. Retinol is converted into retinaldehyde by retinol dehydrogenases, and retinaldehyde is further converted into RA by retinaldehyde dehydrogenases. Once taken up by cells, RA binds to RA receptors, which then interact with RA response elements (RAREs) in the enhancer regions of target genes to activate RA-specific gene transcription [63]. While RA signaling is well characterized, particularly in promoting photoreceptor maturation and survival [74, 73, 191], there has been no comprehensive study on how RA influences the spatial organization and development of other retinal cell types during retinogenesis. This is particularly relevant for understanding the formation of the laminar structure and the regional segregation into the macula and peripheral retina. The macula is distinguished by a higher cone density, especially in the fovea, a 0.35mm diameter region crucial for high-acuity vision [67, 68, 192], while the periphery is predominantly composed of rod photoreceptors. The differential distribution of cones and rods reflects overall retinal maturation and patterning.

Retinal organoid models derived from pluripotent stem cells have become a powerful tool to study retinal development as it is able to recapitulate the major cell types of the *in vivo* retina with high transcriptional similarity [193]. To date, numerous single-cell transcriptomics studies have assessed the fidelity of retinal organoids in recapitulating the major retinal cell types and rare subtypes, against the human retina [194, 195, 195, 196]. However, there has only been a recent emergence of spatial transcriptomics and integration of chromatin accessibility studies to understand how the spatial patterning of the organoid impacts the multi-faceted nature of retinogenesis [197, 198, 199].

Despite the progress in mimicking the growth and maturation of retinal cell types within a 3D organoid structure, current protocols have limitations in fully achieving the regional specification of the retina [66], often resulting in the intermixing of cone and rod photoreceptors in the outer nuclear layer. My aim was to use multi-resolution spatial transcriptomics and scRNA-seq to characterize how varying concentrations of RA influence retinal cell fate and spatial organization in retinal organoids, using these models as a proxy for human retinal development. We perform multi-resolution spatial transcriptomics, combining aggregated spot-level and single-cell level data, alongside single-cell transcriptomics to explore the heterogeneity and spatial patterning elicited by RA signals. I envision that this will lead to a more faithful model upon which to study retinogenesis in health and the pathologies of retinal degeneration and macular-related diseases.

4.3 Materials and methods

4.3.1 Stem cell maintenance and differentiation

Cell lines in this study included A81.BRN3b.mCherry (A81) [200], a H7 Human ESC line reporter line, and 149br.CRX.Cherry human iPSC reporter line (148br) (donated by Robyn Jamieson, CMRI). Cells were incubated at 37 °C 5% CO2. hiPSCs were grown and expanded in feeder-free conditions using Essential 8 media (E8, Thermo Fisher Scientific; #A1517001) on 6 well plates coated with Geltrex (Thermo Fisher Scientific; #A1413302) at a concentration of 1:100. Media was replaced daily, and cells were passaged at 70% confluency via 5-10 minute 37 °C incubation with Versene Solution (0.48 mM) (Thermo Fisher Scientific; #15040066) to detach clumps of cells. Cell clumps were resuspended at a ratio of 1:6-1:12 in E8 with 10 μM ROCK inhibitor (Y-27632 dihydrochloride; STEMCELL Technologies #72307) and seeded in fresh Geltrex-coated 6 well plates. For differentiation, hiPSCs were grown to 90-100% confluency.

4.3.2 Differentiation of hPSCs to retinal organoids

Retinal organoids were differentiated as previously described by Gonzalez-Cordero et al. [45] and West et al. [201] with some modifications. Briefly, at 90-100% confluency (denoted day 0), hiPSC media was replaced with Essential 6 (E6, Thermo Fisher Scientific; #A1516401) daily for two consecutive days. On day 2, E6 media was replaced with pro-neural induction media (PIM, Advanced DMEM/F12; #12634028, 1X N2 supplement; #17502001, 1x L-Glutamine; #25030081, 1X MEM-NEAA; #11140050, 10U/mL Pen-Strep; #15070063 and 0.25μg/mL Amphotericin B; #15290026 [all Thermo Fisher Scientific]). Optic vesicles displaying neuro-retinal epithelium were manually isolated using a needle under EVOS XL microscope (Thermo Fisher Scientific) between day 25-35 and transitioned to 3D suspension culture in low-binding 96 well U-shaped plates (Sigma Aldrich, #CLS7007) in ALT70 (Advanced DMEM/F-12, 2X GlutaMAX Supplement; #35050061, 1X B-27 Supplement minus Vitamin A; #12587001, 10U/mL Pen-Strep and 0.25μg/mL Amphotericin B, All Thermo Fisher Scientific 100 μM taurine [Sigma Aldrich; #T8691] and 10% FBS [Bovogen; SFBS]). At day 70, retinal organoids were transferred into low-binding 24 well plates and supplemented with 1μM all-trans RA (RA, Sigma Aldrich; #R2625) to enhance photoreceptor development. At day 90 and until experimental end-point, media was replaced with ALT90

(Advanced DMEM/F-12, 2X GlutaMAX Supplement, 1X B-27 Supplement minus Vitamin A, 1X N2 Supplement, 7 mM glucose; #G6152), 10U/mL Pen-Strep, 0.25µg/mL Amphotericin B, 1X Lipid Mixture, [all Thermo Fisher Scientific] 100µM taurine, and supplemented with 0.5µM RA. Media was replaced every two days and cells maintained at 37 °C 5% CO2. Different treatments were performed by the addition of exogenous RA (10/1/0µM) or BMS-493 (1µM [STEMCELL Technologies; #73974]) to retinal organoids from day 70 (10 weeks) to the experimental end point.

4.3.3 Sample and library preparation for scRNA-sequencing, and data pre-processing

6-7 organoids differentiated from the A81.BRN3b.mCherry reporter line were collected for each condition in a 1.5 ml tube and washed twice with PBS. Retinal organoids were dissociated into single cell suspension using the Neurosphere Dissociation Kit (P) (Miltenyi Biotec). Enzymatic digestion was performed as per manufacturer protocol for 10 minutes at 37 °C with intermittent agitation, followed by gentle mechanical dissociation, and a further incubation for 5 minutes at 37 °C. The cell suspension was dissociated into single cells by pipetting before the enzymatic reaction was stopped by washing with HBSS. The cell suspension was filtered through MACS SmartStrainer 30µm (Miltenyi Biotec) before being pelleted by centrifugation at 400g for 10 minutes at room temperature. The cell pellet was resuspended in ALT90 and maintained on ice.

Each single cell suspension of dissociated retinal organoids was assessed for viability using 0.4% Trypan Blue staining on a Countess II Automated Cell Counter (Invitrogen) and concentration was adjusted to 1000 cells/µL. Cell suspension was loaded on a single-cell-B Chip (10X Genomics) for a target output of 10,000 cells per sample. Single-cell droplet capture was performed on the Chromium Controller (10x Genomics). cDNA library preparation was performed in accordance with the Single-Cell 3' v3 protocol. Libraries were evaluated for fragment size and concentration using Agilent HSD5000 ScreenTape System. Samples were sequenced on an Illumina NovaSeq6000 instrument according to the manufacturer's instructions (Illumina). Sequencing was carried out using 2×150 paired-end (PE) configuration with a sequencing depth of 40,000 reads per cell. The sequences were processed by GENEWIZ, China. Basecalling is performed by Illumina RTA software in sequencer, further demultiplexing is performed by Illumina bcl2fastq 2.20.

Counts were generated using 10x Genomics Cell Ranger (v7.2.0) [138] and reads were mapped to the GRCh8 genome. Empty droplet filtering was performed using DropletUtils (v1.22.0) [139]. After standard cell filtering, doublet detection was performed using DoubletFinder (v2.0.4) [140].

4.3.4 Sample and library preparation for 10x Visium, and data pre-processing

Differentiated retinal organoids (Table 4.1 following the control protocol were collected at 10 weeks (Visium 10x platform; #PN-1000187) and washed with PBS before cryo-embedding in OCT (Leica; #14020108926) using liquid nitrogen cooled 2-methylbutane (Sigma Aldrich; #M32631) before storage at -80 °C. Frozen organoids were sectioned (10 µm thickness) using a cryostat (Leica; #CM1950) onto 10x Visium library preparation slides and then fixed for 10 minutes at -20 °C using methanol. Slides were stained with DAPI and imaged on the Leica DMI8 Stellaris 8 confocal microscope, using tile scanning to stitch together an image of the whole capture area. Tissue permeabilisation (1.5 minutes), cDNA synthesis, amplification and library preparation were performed according to the Visium Spatial Gene Expression User Guide (CG000239). Slides were sent for sequencing by Azenta, Genewiz using the Illumina HiSeq instrument according to the manufacturer's instructions (Genewiz). Sequencing was carried out using 2×150 paired-end (PE) configuration with a sequencing depth of 30,000 reads/spot. Basecalling is performed by Illumina RTA software in sequencer, further demultiplexing is performed by Illumina bcl2fastq 2.20.

Table 4.1. Number of organoids collected for 10x Visium. Total number of organoids per cell line and replicate slide quoted in the table.

Cell line	Slide 1	Slide 2
149br.CRX.Cherry	12	15
A81.BRN3b.mCherry	7	6

Counts and spatial locations were obtained by performing the standard 10X Space Ranger (v3.0.1) pipeline and reads were mapped to the GRCh8 genome. Spatial counts for individual organoids were generated using a fiducial alignment and tissue assignment JSON file generated in the Loupe Browser on the input brightfield image of each Visium slide.

4.3.5 Sample and library preparation for STOmics Stereo-seq, and data pre-processing

Differentiated retinal organoids after treatment with RA 1µM/0.5 µM (control), RA 10µM, RA- and BMS-493 1µM were collected at day 112 (16 weeks) and day 182/185 (26 weeks) then washed with PBS (Table 4.2). Then, organoids were cryo-embedded in OCT (Leica; #14020108926) using liquid nitrogen cooled 2-methylbutane (Sigma Aldrich; #M32631) before storage at -80 °C. Embedded organoids were sectioned (10µm thickness) using a cryostat (Leica CM1950) onto library preparation chips from the BGI Stereo-seq platform (#111KT114). Chips were fixed for 10 minutes at -20 °C using methanol. Slides were stained with Qubit ssDNA stain (Thermo Fisher Scientific, #Q10212) and then imaged on the Zeiss Axio Imager 2, using tile scanning to stitch together an image of the whole capture area. Tissue permeabilisation (5 minutes), cDNA synthesis, amplification and library preparation were performed according to the Transcriptomics Set For Chip-on-a-slide User Manual (211ST114, Version C). Chips were sent for sequencing by Azenta, Genewiz using the MGI T7 instrument according to the manufacturer's instructions (Genewiz). Sequencing was carried out using 2×100 paired-end (PE) configuration with a sequencing depth of 900M reads/chip. Basecalling is performed by Illumina RTA software in sequencer, further demultiplexing is performed by Illumina bcl2fastq 2.20.

The SAW default pipeline was run to align and map reads to the GRCh8 genome. For each slide, multiple FOVs were manually stitched together into a single panoramic image using the ImageStudio (v3.0.0) software and registered to the gene expression matrix using the SAW registration module (v7.1). Segmentation of the tissue coverage area was performed based on the ssDNA stain.

Table 4.2. Number of organoids collected for BGI Stereo-seq. Total number of organoids per condition for each time point quoted in the table.

Condition	D112 (16 weeks)	D182 (26 weeks)	D185 (26 weeks)
RA 1/0.5µM (control)	15	8	3
BMS-493 1µM	9	4	0
RA 10µM	9	1	3
RA-	21	12	0

4.3.6 Cell type annotation

To identify the closest equivalent time point of *in vivo* retinal development with organoids profiled using 10x Visium at 10 weeks of culture, overall transcriptomic similarity was compared between the organoids and Zuo et al. retinal atlas [202]. Pearson correlation was calculated between each time point represented in the Zuo et al. retinal atlas against the average transcriptomic profiles across all organoids in each 10x Visium slide. Feature selection using highly variable genes was performed prior to calculating the transcriptomic correlation. Cell type deconvolution of the 10x Visium datasets was performed using the CARD algorithm [5] where multiple cell types are represented as a fraction of a single spot.

To annotate the scRNA-seq data, an integrated reference of multiple human retinal studies from the scEiaD ocular meta-atlas [203] (Table 4.3) was created using the default Seurat v5 integration pipeline [41], where the reciprocal PCA method was used. After batch correction, the transfer learning approach was applied to classify cell types.

Stereo-seq data was analysed at bin60 resolution (30μm x 30μm), which is approximately at single-cell resolution. Specifically, matching 16 weeks scRNA-seq data was used as the reference for each condition profiled spatially at 16 weeks. To annotate the 26 weeks spatial data, previous in-house generated retinal organoids at 30 weeks derived from two different cell lines, HPSI0314i-hoik_1 (hoik) UCLOOi017-A-1 (UCL) were collated. Derivation of these retinal organoids is previously described in Kim et al. [193] and Fernando et al. [123]. Additionally, the *in vivo* retinal atlas from Zuo et al. [202] at post-conception 24 weeks was included to provide greater coverage of all cell types and all sources were integrated to form a single reference upon which label transfer using Seurat (v5.0.1) was performed.

4.3.7 Transcriptomic classification of macula and periphery regions

To select the appropriate time points, cell identity similarity of control 16 weeks organoids was compared against each time point in the Zuo et al. retinal atlas [202]. Cell identity statistics were generated Using Cepo [204] and Pearson correlation was calculated between equivalent cell types between the reference and the control organoids. Then, the correlation score across all cell types was averaged to determine the best matching time point with the highest average correlation. Then, using scClassify [40] with default parameters, a HOPACH hierarchical tree was constructed for each reference time point, simultaneously learning the region and cell type information. Using a weighted kNN classifier, cells with with a classification confidence of at least 0.5 were assigned a cell type label, otherwise, they were unassigned. Then, regional classification confidence was aggregated and compared for each cell type. The retinal pigment epithelium (RPE) tissue was removed prior to avoid misclassification as the reference does not contain RPE tissue.

Table 4.3. Summary of retinal studies used for scRNA-seq annotation. Author, accession number, tissue source and ages are quoted in the table. *hgw = human gestational weeks*, *pnd = post-natal days*

Name	**SRA Accession Number**	**Tissue Source**	**Ages**
Lukowski et al.	E-MTAB-7316	*in vivo*	post-mortem unknown
Cowan et al.	EGAD00001006350	*in vivo*/*in vitro*	post-mortem 50-80 years adult, 30/38 weeks organoids
Lu et al. Set 1	SRP223254	*in vivo*/*in vitro*	post-mortem adult, post-mortem hgw 9-19 weeks, 24-59 days organoids
Lu et al. Set 2	SRP170761	*in vivo*	post-mortem 8 pnd
Menon et al.	SRP222958	*in vivo*	post-mortem unknown
Sridhar et al.	SRP238587	*in vivo*/*in vitro*	45-205 days organoids, postmortem 59 - 125 days fetal
Gautam et al.	SRP255012	*in vivo*	post-mortem 24 - 84 years adult

4.4 Results

4.4.1 Timecourse analysis of multi-resolution single-cell and spatial transcriptomics in retinal organoid development

To characterise the initial emergence of RPCs and differentiation into the early-born retinal cell types prior to the commencement of RA modulation, we performed 10x Visium spatial transcriptomic profiling of retinal organoids at 10 weeks (day 70) of growth. Retinal organoids derived from an ESC-derived cell line, A81.BRN3b.mCherry and an iPSC-derived cell line, 149br.CRX.Cherry (Table 4.1) were compared, to assess their differentiation propensities toward the retinal lineage. As RA has been shown to regulate photoreceptor maturation in a dose-dependent and time dependent-manner [205, 70], we tested different concentrations to examine the effect in retinal organoids. From 10 weeks onwards, we exposed the retinal organoid culture to a default RA concentration optimized from Gonzalez-Cordero et al. [45] and West et al. [201], and an extremely high concentration (10 µM). Additionally, I assessed the effects of intrinsic basal RA levels (no RA) and inhibition of RA using BMS-493, a pan-RA receptor inverse agonist [206, 207], at a concentration matching the default RA level. This promotes competitive binding to inhibit RA receptor activity, thus repressing transcription of downstream RA target genes (Figure 4.1).

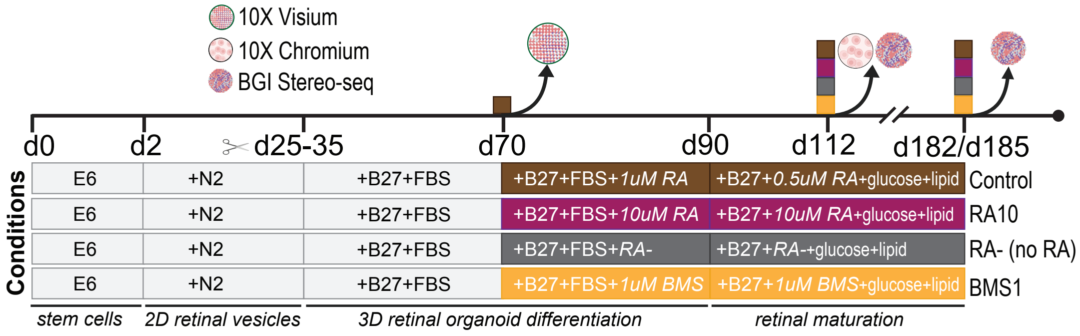


Figure 4.1. Multi-resolution single-cell and spatial transcriptomics profiling of RA perturbation during retinal organoid development. Experimental design schematic showing the differentiation of iPSCs into retinal vesicles in 2D monolayer culture before excision to encourage 3D structural formation into mature retinal organoids, following the previously established protocol [45, 201]. Organoids were treated from day 70 onward with varying concentrations of RA, and the inhibition by BMS-493. 10x Visium was performed at day 70 only, followed by scRNA-sequencing (n=6-7) and BGI Stereo-seq performed at 16 weeks (day 112), and BGI Stereo-seq at 26 weeks (day 182/day 185). Organoid numbers for spatial transcriptomics are listed in Tables 4.1, 4.2.

To uncover the cellular heterogeneity at 16 weeks which marks the beginning of late-phase retinogenesis, we profiled the single-cell transcriptome to characterise the transcriptomic changes as a result of RA perturbation. Additionally, to investigate how the spatial organisation of retinal cell types is affected, we profiled the approximate single-cell spatial transcriptome using the BGI Stereo-seq platform at 16 weeks and 26 weeks of retinal organoid development, where maturation is reached by the experimental end point.

4.4.2 Spatial transcriptomics reveals signatures of early retinogenesis

Prior to RA treatment, I first mapped the spatial distribution of early-born retinal cells at 10 weeks of culture and determined whether the 3D environment of retinal organoids is able to recapitulate the dorsoventral patterning typically observed during human retinal development. I used the *in vivo* human fetal retina atlas compiled from scRNA-seq of human embryos and foetuses aged between 8-24 weeks post-conception as a reference for cell type deconvolution.

To determine the most appropriate *in vivo* time point of the Zuo et al. retinal atlas [202] as a reference for cell deconvolution, I calculated the average of the whole transcriptome correlation using genes that intersect with the top 80% of highly variable genes at each time point compared across all slides and found that 11 post-conception weeks had the highest transcriptional similarity (Figure 4.2 A) that closely matches the *in vitro* organoid age. After performing cell type deconvolution using CARD [5], I visualised the cell type proportions of each spot per cell type of each organoid (Figure 4.2B, C). As expected, the majority of the organoids is dominated by high proportions of RPCs which is consistent across the two cell lines, followed by cones, retinal ganglion cells and horizontal cells. The addition of exogenous RA may optimize for the abundance of photoreceptors, which obeys with what is expected in normal retina development.

When I visualised the spatial pattern of two widely reported dorsal-ventral markers, *ALDH1A1* which defines the dorsal retina and *ALDH1A3*, which defines the ventral retina [63], there is specific spatial localisation of the gene expression in both cell lines (Figure 4.3A), however stronger levels of expression is observed in the 149br cell line-generated organoids. This suggests that the organoids may possess the capacity to establish the natural orientation of the retina along the dorsoventral axis.

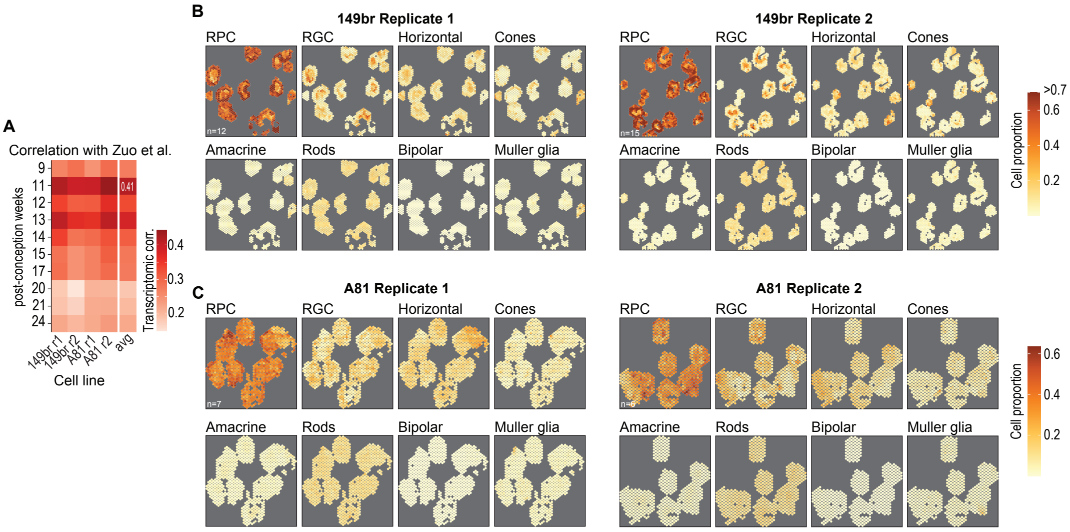


Figure 4.2. Spatial transcriptomics reveals the presence of early-born retinal cell types and dorsal-ventral patterning in the retinal organoid. (a) Comparison of averaged transcriptomic profiles of organoids in each 10x Visium against the whole transcriptome of each time point in Zuo et al. [202]. Transcriptomic similarity is measured by average Pearson correlation across all slides per time point. (b,c) Cell proportions estimated per cell type in each organoid derived from the 149br and A81 cell lines, respectively. Cell types are ordered by developmental timing during retinogenesis.

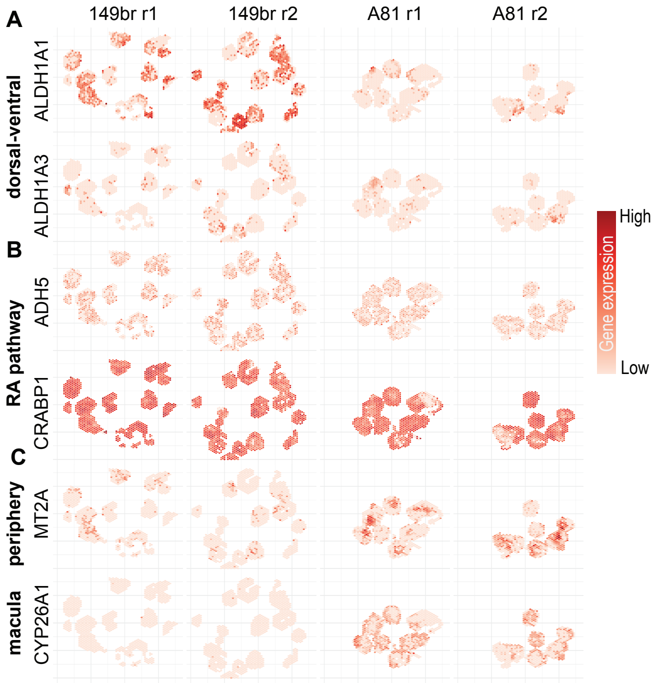


Figure 4.3. Spatial gene expression reveals dorsoventral patterning of the retinal organoids. (A-C) Log-normalised gene expression of select markers defining the (A) dorsoventral axis, (B) RA pathway and (C) macula and periphery.

I additionally visualised two genes *CRABP1* and *ADH5* [190] (Figure 4.3B) to show that RA is actively metabolised. Finally, I selected an example of genes posited to be differentially expressed in the macula and periphery regions of the retina. *CYP26A1* is a RA metabolizing enzyme, which has been shown to be differentially expressed in the foveal retina [208, 209] and *MT2A* is a putative periphery retina marker [209].

The A81 cell line-generated organoids showed higher spatial expression of these genes on the border and central region of the organoid, respectively. However, only minimal expression of these genes was observed in the 149br cell line-generated organoids, highlighting the discrepancy between organoids generated from two different cell lines. Together, these results infer that the organoids derived from the A81 cell line may have established tissue patterning to an extent, which is associated with the RA signalling activity. In view of that regionalisation of the macula and periphery follows after dorsoventral patterning, this may suggest that the A81 ESC cell line has a greater differentiation propensity toward the retinal lineage compared to the 149br iPSC cell line or the timing of differentiation proceeds at a faster rate. Hence, all subsequent experiments were conducted with the A81 cell line.

4.4.3 Single-cell transcriptomic characterisation of RA perturbation

I sought to characterise the retinal organoid culture at 16 weeks to examine the effect of RA modulation on the formation of retinal cell types, hence inferring the maturity of the organoids based on its progression along retinogenesis. As in-house data was unavailable to serve as a reference for single-cell annotation, I integrated various public sources of retinal cell atlases derived from *in vivo* tissue and *in vitro* organoids of different ages (Figure 4.4A, Table 4.3) collated from the wider scEiaD eye atlas [203]. A previous benchmarking of retinal organoid protocols [193] have shown that the cell identity of retinal cell types is similar between mature and fetal tissue, and most existing retinal organoid protocols can recapitulate the cell identity signatures with high fidelity. Given this, I included multiple sources to increase the coverage and power to detect retinal cell types with relatively lower abundances in our single-cell data. All major cell types of the retina were identified under all conditions (Figure 4.4 B), including the layer of retinal pigment epithelium (RPE) adjacent to the peripheral retina, which has not been found in organoids generated by current protocols [193].

By observing the total cell counts per condition, the control, RA 10µm and RA- conditions showed similar cell numbers, and the BMS 1µm condition may have led to greater cell death as indicated by the low total cell count relative to other treatment conditions (Figure 4.4C). There were, however, different proportions of RPCs, bipolar cells, rods, cones and Müller glia in the organoids across the treatments (Figure 4.4C). Notably, the RA- and BMS 1 μm conditions have lower proportions of RPCs and Müller glia, accompanied by an increase in proportions of bipolar and cone cells. The inhibition of RA signalling activity led to a reduction in rod photoreceptors whereas the lack of RA supplementation led to a slight increase in the proportion of rods.

To assess regional transcriptional identity, the Zuo et al. retinal atlas [202], which profiles macula and peripheral regions from 11 to 24 weeks post-conception, was used as a reference. Cepo was applied to [204] to quantify the extent a gene contributes to the cell identity of a given cell type. This was performed separately for the Zuo et al. reference and the control treatment, followed by correlating the gene-level statistics for matching cell types to derive an aggregated measure of cell identity (Figure 4.4D). I averaged all cell identity scores across each time point, where the highest aggregated score represents the closest transcriptional similarity to internal data. The results showed that the 16 week organoids were transcriptionally similar to retinal ages 12-24 post-conception weeks. Cowan et al. [196] previously applied a linear regression model to predict the retina-equivalent ages of organoid samples and demonstrated that the organoid age closely mimics the timeline of *in vivo* retinal development. Hence, despite the similar magnitude in the average cell identity correlations, the closely corresponding ages 15, 20, 21 weeks were selected for subsequent analysis.

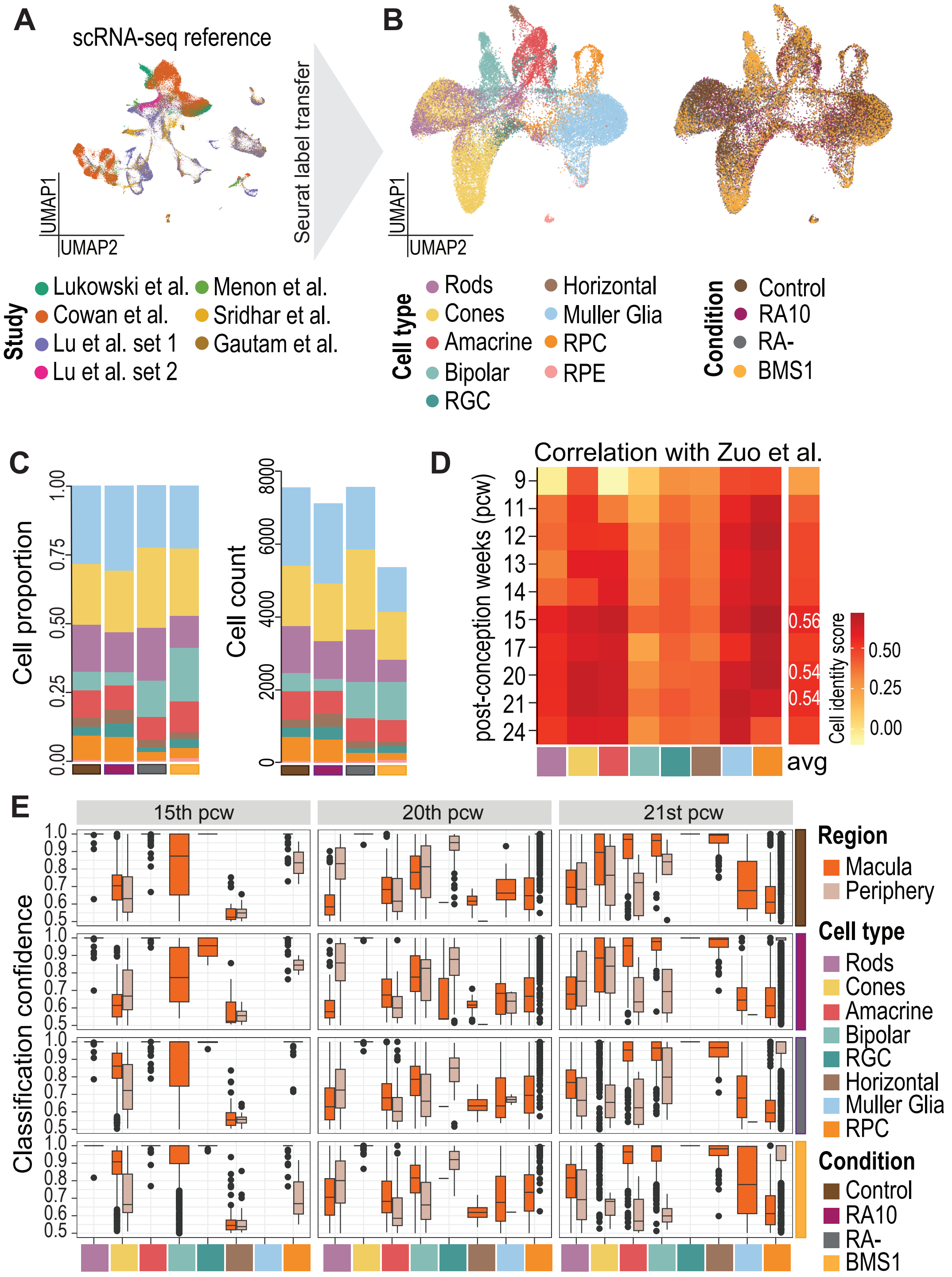


Figure 4.4. Modulation of RA alters retinal cell abundance. (A) Integrated scRNA-seq reference used to annotate each condition at 16 weeks. Details are listed in Table 4.3. (B) Integrated UMAP coloured by cell type (left) and treatment condition (right). (C) Cell type proportions (left) and total cell count (right) per treatment condition. (D) Cell identity scores compared between matching cell types found in the Zuo et al. retinal atlas [202] for each time point against the control treatment condition. The average cell identity score is calculated by summarising the cell identity scores for each time point across all cell types. (E) Classification confidence of scClassify annotations for each cell type, separated by region and each treatment condition. Respective time points shown were used as the reference to construct the HOPACH hierarchical tree.

For these three time points, I used HOPACH to construct a hierarchical tree for each reference time point, simultaneously learning the region and cell type hierarchy using scClassify [40]. For each treatment condition, I predicted the corresponding region and cell type using a weighted kNN classifier (Figure 4.4E). Here, I show the classification confidence to assess whether a cell can be confidently classified two-fold as belonging to a specific region and a specific cell type. Broadly, not all cell types were successfully classified to a specific region. However, regardless of time point, the amacrine, bipolar and cone cells have a stronger macula identity relative to the periphery. At 21 weeks, the highest classification confidence was associated a majority of cells being assigned to a region, indicating that 21 weeks is best equivalent to our *in vitro* 16 weeks data. While we observed a macula bias for some cell types, there was no dominant trend across the different conditions. This suggests that it is not feasible to infer the regional identity based on the single-cell transcriptome and it is necessary to validate the positional information of the cell types in the macula and periphery regions of the organoid.

4.4.4 Spatial heterogeneity of retinal organoid populations during mid to late retinal development under RA modulation

Next, we generated the spatial transcriptome of developing retinal organoids treated with different concentrations of RA at 16 weeks and 26 weeks to examine the spatial organisation of retinal cell types during late retinogenesis. The scRNA-seq data of organoids at 16 weeks under the same treatment conditions was used as the reference to annotate cell types in the Stereo-seq dataset generated at 16 weeks (Figure 4.5A). All cell types display spatial patterning. Specifically, there was a general radial distribution of late-born retinal interneurons such as amacrine and Müller glia cells in the central region of the organoid versus the relatively early-born retinal cell types such as cones and horizontal cells respectively to the peripheral borders of the organoid. The RPE layer where detected, is found at the border of the organoid.

In-house retinal organoid protocols cultured until 30 weeks derived from two different cell lines as previously described [193, 123], and the 24 weeks post-conception time point from Zuo et al. [202] was integrated to increase coverage of rare cell types (Figure 4.5B) for annotation of the Stereo-seq data generated at 26 weeks (Figure 4.5C). RPE was not generated in the in-house protocols or present in the public references, hence we limited the inferences to the retina from hereon. By 26 weeks, there is a dominating presence of Müller glia cells seen in all conditions and comparatively fewer or absence of other retinal interneurons. Comparing the cell type proportions across different conditions in individual organoids (Figure 4.5D) revealed large variations for most cell types across all conditions. The proportion of RPCs from 16 weeks to 26 weeks decreased in all conditions, suggesting that RPCs have ceased expansion. At 16 weeks, cone and rod photoreceptors reached a peak proportion in the BMS 1 μm condition subsequently decreased by half-fold at 26 weeks suggesting the inhibition of RA signalling has impacted photoreceptor survival. In contrast, the RA- treated organoids at 26 weeks contained a higher proportion of photoreceptors. Consistent with the spatial visualisation, there is overall greater proportions of Müller glia at the expense of other retinal interneurons such as bipolar and amacrine cells, and retinal ganglion cells.

Together, the progression of RA perturbation at 16 weeks and 26 weeks shows that RA affects the maturation of the retinal organoid, leading to changes in abundance of early-born and late-born retinal cell types. With the addition of the spatial dimension, I observed spatial patterning of cell types which may be reflective of retinal laminae in the retinal organoids.

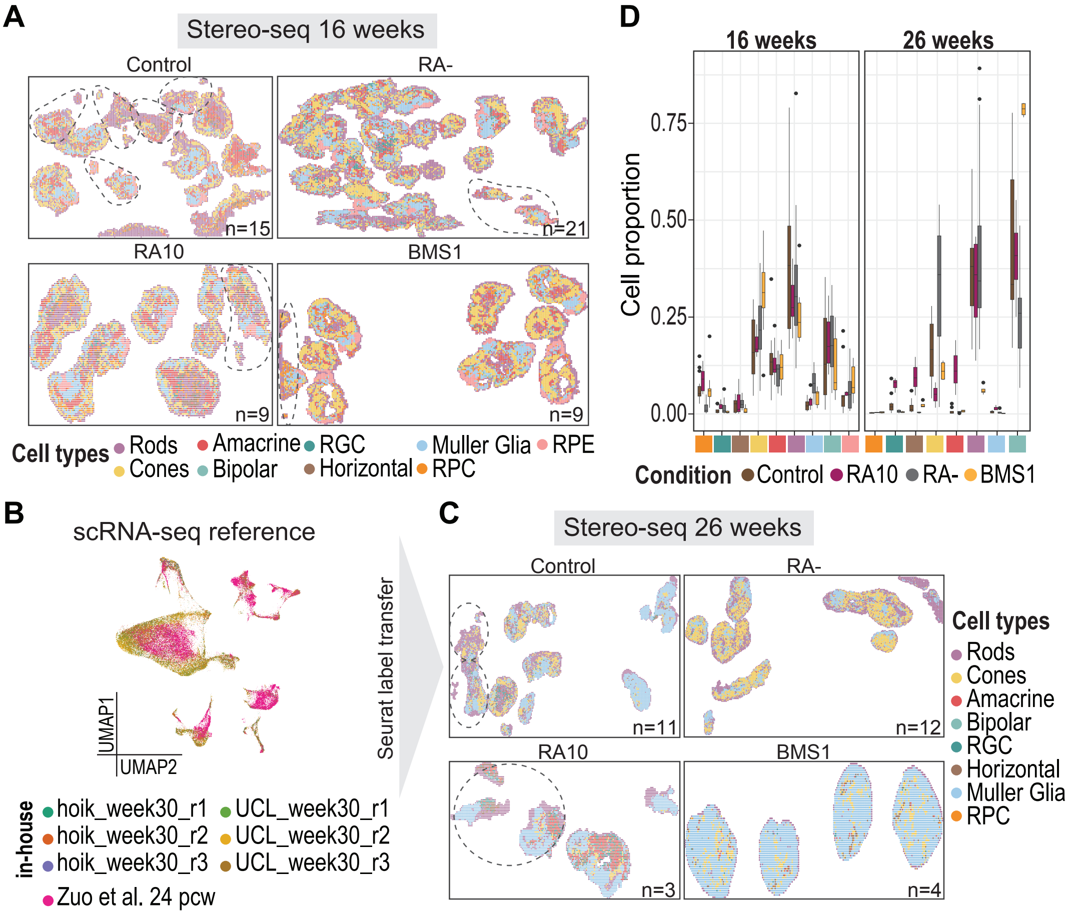


Figure 4.5. Spatial segregation of retinal cell types reflects gradient of organoid maturation. (A) Spatial visualisation of cell types in the 16 weeks retinal organoids separated by RA perturbation, profiled by the BGI Stereo-seq platform. Each spot detected is equivalent to bin60 (30µm x 30µm diameter). Organoid numbers per treatment condition are quoted in Table 4.1. Fragments belonging to a single organoid are encircled in grey for clarity. (B) UMAP visualisation of in-house and public references integrated to form the scRNA-seq reference for annotation of 26 weeks organoids profiled by the BGI Stereo-seq platform. (C) Spatial visualisation of cell types in the 26 weeks retinal organoids separated by RA perturbation, profiled by the BGI Stereo-seq platform. Each spot detected is equivalent to bin60 (30µm x 30µm diameter). Organoid numbers per treatment condition are quoted in Table 4.2. Fragments belonging to a single organoid are encircled in grey for clarity. (D) Boxplots of cell proportions for each cell type. x-axis represents each cell type ordered by development timing during retinogenesis, y-axis represents the cell numbers proportional to the total spots per organoid. Boxplots are coloured by condition. For all boxplots: centre line, median; box limits, upper and lower quartiles; whiskers, 1.5 times the interquartile range; points, outliers.

4.5 Discussion

There are numerous existing scRNA-seq studies that characterize retinal organoid development [195, 208, 196, 194], focusing on the capability of current protocols to replicate the *in vivo* retinal architecture and uncover retinal heterogeneity. Emerging studies in the field have explored the simultaneous profiling of chromatin accessibility, single-cell transcriptomics and low resolution spatial proteomics or spatial transcriptomics [199, 197] to collectively characterise the gene regulatory activity across time in a spatial context. However, there is a lack of comprehensive, high-resolution spatial mapping studies integrated across developmental time and moreover, current studies have yet to perform an in-depth study of how RA, a key signalling pathway, affects retinal cell maturation and spatial organization. In this study, I address these gaps by presenting a multi-resolution, whole-transcriptome spatio-temporal map of retinal organoid development. This approach also explores how RA signaling modulates cell fate and spatial structure, with the goal of optimizing organoid protocols to more accurately mimic the regional organization of the human retina with a central and periphery region.

Previous protocols [70, 45, 201, 210] commonly incorporate exogenous RA to promote photoreceptor maturation, so I investigated the effect of endogenous RA signaling to determine whether exogenous RA truly improves outcomes to generate high fidelity retinal organoids. It is well-known that all-trans-RA promotes the differentiation of cone and rod photoreceptors independently of cell proliferation [191, 70, 73, 205]. Consistent with observations from previous studies that add exogenous RA, I found that adding $0.5\mu M/1\mu M$ RA increases the proportions of rod photoreceptors. However, I also found that excessive RA can trigger a negative feedback loop, which degrades RA and thus leads to a reduction in rod numbers compared to RA-free conditions. This behavior was inversely mirrored in cones, though the mechanism remain unclear, but I propose that cone survival may be linked to RA produced by neighbouring Müller glia [74], as suggested by the correlation between low RA conditions and a high cone-to-glia ratio.

The spatial transcriptomics analysis at 10 weeks revealed that the A81 ESC-derived cell line was more amenable to retinal differentiation compared to the 149br iPSC-derived cell line. Typically, retinal organoid generation relies on the successful differentiation of either iPSC- or ESC-derived cell lines [70, 45, 195]. The pioneering work by Takahashi and

Yamanaka [211, 121] demonstrated that the combined over-expression of Oct4, Sox2, c-Myc and Klf4 is capable of reprogramming fibroblasts to ESC-like cells, effectively rewiring the somatic epigenetics to permit the activation of the pluripotency network. However, improper reorganization of chromatin architecture during this reprogramming process can present a barrier to fully establishing the epigenetic landscape required for ESC-like pluripotency [212]. This epigenetic instability may explain why the 149br iPSC-derived cell line exhibited a delayed differentiation timing compared to the A81 ESC-derived line, which may retain a more stable pluripotent state conducive to retinal differentiation.

Given the sequential differentiation of retinal cells and the formation of retinal laminae, the ratio of early- and late-born cell types offers insight into regional segregation into macular and peripheral regions. The investigation of retinal organoids at 16 weeks revealed that RA-free conditions better preserved photoreceptor survival and maintained proportions of Müller glia. I also found that RA alters the proportions of amacrine, bipolar and ganglion cells which was previously unreported. However, at 26 weeks, an unexpected overabundance of Müller glia across all conditions may suggest cell death of other retinal cell types due to the build-up of a necrotic core [213] but confirmation of apoptotic marker expression is needed. Nevertheless, the spatial patterning observed likely reflects the radial maturation of retinal cells, a process that has also been observed in immunohistochemistry and in situ FISH studies [199], with cone-dense areas resembling the formation of the macula-like region.

While these analyses uncovered the cellular heterogeneity of retinal organoids, the technological and methodological approaches are subject to a few limitations. First, analysis of early retinogenesis in week 10 retinal organoids was found to have a high proportion of RPCs, RGCs, horizontal cells and photoreceptors, with very few spots showing low proportions of cell types typically associated with late retinogenesis, as observed in *in vivo* development. Due to the limited resolution of the 10x Visium technology, where each bead captures the gene expression of multiple cells, I employed the cell type deconvolution algorithm, CARD [5], to estimate the proportions of different cell types within a single spot. However, a limitation of such algorithms, which rely on non-negative matrix factorization, is the inherent assumption that the observed data can be fully explained by a linear combination of cell types present in the scRNA-seq reference, without accounting for any error term for unassigned cell types or confidence measure of the cell type assignment. Although it is possible that the temporal progression of *in vitro* protocols differs from normal retinal development, this limitation of the algorithm implies that subsequent inferences are highly dependent on choosing the appropriate reference dataset. In Figure 4.2A, Pearson's correlation between top 80% HVGs was performed to select the optimal corresponding *in vivo* timepoint as a reference to annotate the 10x Visium data. Although 11 weeks displayed the highest magnitude of average correlation, the overall modest correlation values (< 0.5) could reflect the insufficient resolution of additional information such as cell-type specific or maturation-specific genes to refine the gene set to derive more biologically meaningful correlations. These reasons may explain why low proportions of cells that typically arise during late retinogenesis were detected, which does not reflect the timing of *in vivo* development.

Similarly, at 16 weeks the average cell identity scores are similar across most time points. As retinal organoids at this stage of development are characterised by the continued diversification of retinal progenitor cells to early- and late-stage retinal interneurons which may explain the observed similarity in overall transcriptional profiles between early and late post-conception weeks, reflecting an intermediary phase with a high proportion of differentiating cells. This is consistent with findings in Cowan et al. [196], where the pairwise correlations with the human retina and organoid samples show that 18-week organoids correlate similarly with approximate retina ages from 15-20 weeks. Again, an alternative approach could involve the derivation and Cepo correlation of retinal maturation genes rather than the whole transcriptome. This may provide more biologically-informative correlations and better measure the extent of retinal maturation in the organoid relative to the retinal tissue.

In addition, the Stereo-seq data are not true single-cell resolution as the ssDNA stain for nuclei precludes the segmentation of individual cells without multiplex immunofluorescence staining to highlight the cell boundary. Hence, total counts per square area were aggregated into bins to approximate single-cell resolution for spatial analysis of Stereo-seq. As highlighted in a recent systematic comparison of sequencing-based spatial transcriptomics technologies [214], the limited sequencing depth of the Stereo-seq platform may have compromised the sensitivity to detect rare cell types such as retinal interneurons where there are very low cell proportions.

The differences between cell type proportions seen between the scRNA-seq and Stereo-seq platforms may also be explained by the challenges during sample preparation. Since scRNA-seq relies on dissociating organoids into single cells, certain cell types may be more challenging to dissociate, potentially impacting the representation of certain cell types and limit their downstream detection. On the other hand, as the quality of spatial transcriptomics is highly dependent on the tissue quality of individual slices, the accumulation of necrosis in organoids would be susceptible to this limitation which likely compromised the data quality. Unlike scRNA-seq, which includes steps to remove dead cells and enrich for viable cells, such measures are not feasible for spatial approaches, which may contribute to discrepancies in cell type estimates between the two technologies. Additionally, scRNA-seq involves pooling multiple organoids, offering a more comprehensive cellular representation, whereas spatial transcriptomics is constrained to single organoid slices, introducing variability depending on the cross-section analysed.

Overall, the study sets a precedent for multi-modal profiling of retinal organoid cultures to comprehensively characterize the effects of RA perturbation on retinal cell fate. Specifically, I emphasize the importance of spatial transcriptomics and spatial proteomics in evaluating how well current protocols emulate the spatial organization of *in vivo* retinal structures, where axes and regions are defined by the specific spatial organisation of the retinal cell types responsible for different aspects of vision.

4.6 Summary

In chapter 4, I explored changes in cellular heterogeneity and spatial organisation of retinal organoids undergoing retinogenesis across time from the beginning of retinal cell fate commitment to late development under RA perturbation. To date, retinal organoid studies mostly profile the single-cell transcriptomics, or low-resolution spatial transcriptomics which limits the inference of accurate cellular heterogeneity. Therefore, I addressed this gap in the field by presenting a multi-resolution spatial transcriptomics study with paired single-cell transcriptomics to integratively infer and characterise changes during retinogenesis under perturbation to study how changes in RA signalling influences retinal cell fate. Under the current results, intrinsic endogenous RA signalling is sufficient to maintain high proportion of rod and cone photoreceptors relative to other treatments. In areas where there is high cone density could be indicative of a macula-like region. However, additional experiments to quantify the rod/cone density through imaging analysis of the retinal laminae is required for sufficient evidence of a macula-like region.

CHAPTER 5

Conclusions and future directions

5.1 Summary of the thesis

In this thesis, I focus on applying computational systems approaches to investigate the foundations of spatial and multi-omics analyses of molecular programs and regulatory networks that underlie cell identity and cell fate decisions, and lineage commitment during early development using 3D organoids as an *in vitro* model. By leveraging spatial transcriptomics, single-cell transcriptomics and bulk multi-omics to profile tissues and organoid models and during their development, these works highlight the importance of a multi-faceted perspective to characterise the coordination of cell signalling that leads to cell identity and lineage commitment during *in vivo* human development. Collectively, these studies demonstrate the central role computational methods play in omic data analyses, and the value of spatial and multi-omics measurements in contextualising the localisation and regulation of gene expression, enabling novel biological insights into the cell phenotype.

Through the evaluation of SVG detection methods, we learned the capability of current computational methodologies to capture the spatial relationship between cells to prioritise genes with variable patterns across space. To this end, we observed that many methods encode spatial relationships using weighted covariance matrices or nearest-neighbor graphs, thereby defining spatial neighborhoods based on the assumption that nearby cells will exhibit similar gene expression patterns. However, the performance of these approaches can be influenced by the sparsity inherent in certain spatial technologies, which impacts the grouping of spatial neighborhoods. Additionally, the variability in mRNA abundance poses challenges when identifying regions with low expression, especially due to high dropout rates typical of high-throughput sequencing technologies. This may provide an explanation as to why most SVG detection methods tend to prioritize highly expressed SVGs. While spatial transcriptomics has advanced the field by allowing the inference of cell-cell interactions and signaling between cell types and visualise how spatial proximity is associated with cellular behaviour, there is still a lack of systematic evaluation of the extent gene expression variation is truly explained by spatial variation.

Moving beyond single modality measurement of transcriptomics, I expanded the analyses to integrate multi-omics using 3D organoid systems given its importance in unravelling the complexity behind developmental systems across multiple regulatory programs. Early human development is a highly dynamic process where the cell undergoes immense growth and increased multi-cellular diversity. During this span of time, it is imperative that correct signalling and gene regulatory programs are established to ensure healthy development. To understand the signalling networks crucial to establish neural cell fate, we employed a human cerebral organoid model to study the temporal dynamics from the exit of pluripotency to the formation of the neuroepithelium. We profiled the phosphoproteome, proteome and transcriptome to characterise how the phosphorylation of kinase substrates propagates further downstream gene transcription and protein translation of key molecular factors that influence the neural transcription program. Alongside the generation of a multi-omics reference cerebral organoid formation and the elucidation of its global dynamics, this work also created future opportunities to investigate the molecular mechanisms of kinase phosphorylation on potentially crucial substrates whose dysregulation has been associated with neurodevelopmental disorders.

Development is also a highly spatially organised process, demanding the proper tissue patterning and morphogenesis of organ structures. To this end, I explored the implication of retinoic acid signalling, a key morphogen on the spatial organisation during retinogenesis through human cell-line derived retinal organoids. Through the multi-resolution spatial transcriptomic and scRNA-seq profiling of varying retinoic acid treatment, I investigated changes in cellular heterogeneity and maturation as a consequence of perturbed retinoic acid signalling activity. The spatial contextualisation paired with gene expression measurement allowed us to observe changes in the spatial organisation of retinal cell types and regional delineation of the central and periphery regions typical of real human retina development. It enabled us to link the changing spatial localisation of retinal cell types directly with transcriptional changes throughout time which is extremely valuable considering retinogenesis is a stringent spatially organised process to ensure the establishment of proper vision and eye function.

Altogether, the analysis and integration of spatial and multi-omics data provide a more comprehensive view of signaling pathways, cellular heterogeneity, and transcriptional activity, which play a central role in lineage commitment during development. The ability to introduce perturbations in these models not only deepens our understanding of fundamental developmental processes but also holds immense translational potential for disease modeling and therapeutic development.

5.2 Future directions

Building on the research conducted in this thesis, I foresee several areas for advancement in spatial biology and offer recommendations for future applications of high-throughput spatial multi-omics technologies in organoid modelling.

The evaluation of current SVG detection methods revealed a high dependency between spatial variance and gene expression level. While the mean-variance relationship is a well-documented bias in high throughput sequencing technologies, analogous highly variable gene detection methods in scRNA-seq account for this bias to retain lowly expressing genes that display transcriptional variability. Therefore, further investigation should focus on distinguishing whether this trend is biologically-motivated or a technical artifact. Furthermore, given biologically meaningful genes may inherently be cell type specific, the field would also benefit from the detection of cell type specific spatially variable genes which would reveal transcriptional signatures across space that together, define cell states or identities.

Although multi-omics profiling of early neural development in cerebral organoids allowed us to identify the phospho-signalling and transcriptional networks active at a global scale that dictates neural lineage commitment, further experimentation is required to connect the kinase activity to causation. I have performed computational prediction of potentially novel AKT substrates and demonstrated in an *in vitro* kinase assay that the AKT kinase indeed phosphorylates the substrate peptides. However, this assay does not reflect real *in vivo* or *in vitro* settings during organoid development nor show whether there is a quantifiable effect on neural development. Therefore, additional evidence is required such as performing perturbation screens [215] to assess the impact of these molecules on the phenotype during differentiation. We note that ultimately, human brain development involves distinct regional segregation wherein each area is responsible for distinct tasks such as sensory perception, and motor and cognitive functions. As was seen in the scRNA-sequencing analysis, cellular heterogeneity in neuronal populations and neural crest cells could eventually lead to the regionalisation in the cerebral organoids and therefore examining the spatial organisation of precursor populations could reveal spatial cues during early neural development. However, as bulk technologies are limited to sample-level inferences, this presents challenges when attempting to integrate spatial technologies with bulk multi-omics. Future tool development may focus on the deconvolution of bulk expression profiles prior to mapping putative cell type populations onto spatial data.

To assess the changes in retinogenesis which are often characterised by differential cell abundance across retinal laminae, we performed spatial transcriptomics at the major phases of organoid growth to examine the impact of retinoic acid on retinal cell fate. Through this, we gained insight into the cellular heterogeneity and spatial organisation of cell types in the developing retinal organoid. While I have focused primarily on exploring the transcriptional dynamics of key molecules in the RA signaling pathway, the mechanism by which RA receptors bind to enhancer elements upstream of their target genes to initiate transcription is well known. Future work would benefit from profiling chromatin accessibility to connect open chromatin regions with the transcription of putative retinoic acid target genes, providing a more comprehensive view of RA-mediated gene regulation. Furthermore, as an external validation of retinoic acid levels, metabolomic profiling could be performed to complement the retinoic acid transcription, as well as the activity of other morphogens such as FGF8 [216, 69] during retinogenesis. In addition, imaging analysis could be performed to study the morphology and proportions of photoreceptors in the retinal laminae of the organoid to infer the distinction between macula and periphery regions in the organoid.

In these two studies, I have demonstrated the potential of using organoids as an in vitro 3D model to study early development and identify key pathways involved in this process. However, organoids are yet to be truly an accurate representation of the tissue or perturbation state that they are trying to model [50]. While advancements in RNA sequencing, proteomics, and spatial transcriptomics now allow for in-depth comparisons of organoids to *in vivo* tissues, integrating multiple modalities over a developmental timecourse would be informative for determining whether organoids truly replicate cellular states accurately, especially in unguided protocols wherein there is more likely to be greater variability in the absence of molecules which promote targeted pathways during development. Assessment of organoid reproducibility [44, 217, 58, 210] will also be crucial to ensure protocols are reliable and that all sources of technical variation in the growth of the organoid are mitigated.

In summary, I have demonstrated the power of computational systems approaches in analyzing spatial transcriptomics and multi-omics data. These methods were applied to characterize tissues and trans-regulatory networks (TRNs) involved in human development using organoid models. Future advancements in computational methodologies should focus on more comprehensive integration of diverse omics data with spatial resolution. This will enable a more holistic understanding of spatially-informed TRNs and their role in shaping cellular phenotypes.

Bibliography

[1] Vivien Marx. 'Method of the Year: spatially resolved transcriptomics'. In: *Nature Methods* 18.1 (2021), pp. 9–14.

[2] Sheel Shah et al. 'seqFISH Accurately Detects Transcripts in Single Cells and Reveals Robust Spatial Organization in the Hippocampus'. In: *Neuron* 94.4 (2017), 752–758.e1.

[3] Chee-Huat Linus Eng et al. 'Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+'. In: *Nature* 568.7751 (2019), pp. 235–239.

[4] Chenglong Xia et al. 'Multiplexed detection of RNA using MERFISH and branched DNA amplification'. In: *Scientific Reports* 9.1 (2019), p. 7721.

[5] Ying Ma and Xiang Zhou. 'Spatially informed cell-type deconvolution for spatial transcriptomics'. In: *Nature Biotechnology* 40.9 (2022), pp. 1349–1359.

[6] Dylan M. Cable et al. 'Robust decomposition of cell type mixtures in spatial transcriptomics'. In: *Nature Biotechnology* 40.4 (2022), pp. 517–526.

[7] Ao Chen et al. 'Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays'. en. In: *Cell* 185.10 (2022), 1777–1792.e21.

[8] Sabrina M. Lewis et al. 'Spatial omics and multiplexed imaging to explore cancer biology'. In: *Nature Methods* 18.9 (2021), pp. 997–1012.

[9] Nicole Cloonan et al. 'Stem cell transcriptome profiling via massive-scale mRNA sequencing'. In: *Nature Methods* 5 (2008), pp. 613–619.

[10] Ruedi Aebersold and Matthias Mann. 'Mass spectrometry-based proteomics'. In: *Nature* 422.6928 (2003), pp. 198–207.

[11] Jesper V. Olsen et al. 'Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks'. In: *Cell* 127.3 (2006), pp. 635–648.

[12] Sean J. Humphrey, David E. James and Matthias Mann. 'Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation'. In: *Trends in endocrinology and metabolism: TEM* 26.12 (2015), pp. 676–687.

[13] Sunil U. Bajad et al. 'Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry'. In: *Journal of Chromatography A* 1125.1 (2006), pp. 76–88.

[14] Caroline H. Johnson, Julijana Ivanisevic and Gary Siuzdak. 'Metabolomics: beyond biomarkers and towards mechanisms'. en. In: *Nature Reviews Molecular Cell Biology* 17.7 (2016), pp. 451–459.

[15] David S. Johnson et al. 'Genome-wide mapping of in vivo protein-DNA interactions'. In: *Science (New York, N.Y.)* 5830 (2007), pp. 1497–1502.

[16] Fiorella C. Grandi et al. 'Chromatin accessibility profiling by ATAC-seq'. In: *Nature Protocols* 17.6 (2022), pp. 1518–1552.

[17] Amy Webb et al. 'RNA sequencing of transcriptomes in human brain regions: protein-coding and non-coding RNAs, isoforms and alleles'. In: *BMC Genomics* 16.1 (2015), p. 990.

[18] Young Mok Park et al. 'Profiling human brain proteome by multi-dimensional separations coupled with MS'. In: *Proteomics* 6.18 (2006), pp. 4978–4986.

[19] Eric B. Dammer et al. 'Quantitative phosphoproteomics of Alzheimer's disease reveals cross-talk between kinases and small heat shock proteins'. In: *PROTEOMICS* 15.2-3 (2015), pp. 508–519.

[20] Mathias Uhlén et al. 'Tissue-based map of the human proteome'. In: *Science* 347.6220 (2015), p. 1260419.

[21] Michael J. Hawrylycz et al. 'An anatomically comprehensive atlas of the adult human brain transcriptome'. In: *Nature* 489.7416 (2012), pp. 391–399.

[22] Katsuyuki Yugi et al. 'Trans-Omics: How To Reconstruct Biochemical Networks Across Multiple Omic Layers'. In: *Trends in Biotechnology* 34.4 (2016), pp. 276–290.

[23] Peter V. Hornbeck et al. 'PhosphoSitePlus, 2014: mutations, PTMs and recalibrations'. In: *Nucleic Acids Research* 43.D1 (2015), pp. D512–D520.

[24] Heonjong Han et al. 'TRRUST: a reference database of human transcriptional regulatory interactions'. In: *Scientific Reports* 5.1 (2015), p. 11432.

[25] Jürgen Cox et al. 'Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ'. In: *Molecular & Cellular Proteomics : MCP* 13.9 (2014), pp. 2513–2526.

[26] Navin Rauniyar and John R. Yates. 'Isobaric labeling-based relative quantification in shotgun proteomics'. In: *Journal of Proteome Research* 13.12 (2014).

[27] Jiaming Li et al. 'TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples'. In: *Nature Methods* 17.4 (2020), pp. 399–404.

[28] Fatima Ardito et al. 'The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review)'. In: *International Journal of Molecular Medicine* 40.2 (2017), pp. 271–280.

[29] Brendan D. Manning and Lewis C. Cantley. ‘AKT/PKB Signaling: Navigating Downstream’. In: *Cell* 129.7 (2007), pp. 1261–1274.

[30] Rachael M. Easton et al. ‘Role for Akt3/Protein Kinase B in Attainment of Normal Brain Size’. In: *Molecular and Cellular Biology* 25.5 (2005), pp. 1869–1878.

[31] Long Wang et al. ‘Brain Development and Akt Signaling: the Crossroads of Signaling Pathway and Neurodevelopmental Diseases’. In: *Journal of Molecular Neuroscience* 61.3 (2017), pp. 379–384.

[32] Méghane Sittewelle and Anne H. Monsoro-Burq. ‘AKT signaling displays multifaceted functions in neural crest development’. In: *Developmental Biology* 444 (Dec. 2018), S144–S155.

[33] Pengyi Yang et al. ‘Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency’. In: *Cell Systems* 8.5 (2019), 427–445.e10.

[34] Hani Jieun Kim et al. ‘Transcriptional network dynamics during the progression of pluripotency revealed by integrative statistical learning’. In: *Nucleic Acids Research* 48.4 (Feb. 2020), pp. 1828–1842.

[35] Suzan Stelloo et al. ‘Deciphering lineage specification during early embryogenesis in mouse gastruloids using multilayered proteomics’. In: *Cell Stem Cell* 31.7 (2024), 1072–1090.e8.

[36] Di Xiao et al. ‘Time-resolved phosphoproteome and proteome analysis reveals kinase signaling on master transcription factors during myogenesis’. In: *iScience* 6 (2022), p. 104489.

[37] Johannes Krumm et al. ‘High temporal resolution proteome and phosphoproteome profiling of stem cell-derived hepatocyte development’. In: *Cell Reports* 13 (2022), p. 110604.

[38] Fuchou Tang et al. ‘mRNA-Seq whole-transcriptome analysis of a single cell’. In: *Nature Methods* 6.5 (2009), pp. 377–382.

[39] Valentine Svensson, Roser Vento-Tormo and Sarah A. Teichmann. ‘Exponential scaling of single-cell RNA-seq in the past decade’. In: *Nature Protocols* 13.4 (2018), pp. 599–604.

[40] Yingxin Lin et al. ‘scClassify: sample size estimation and multiscale classification of cells using single and multiple reference’. In: *Molecular Systems Biology* 16.6 (2020), e9389.

[41] Yuhan Hao et al. ‘Dictionary learning for integrative, multimodal and scalable single-cell analysis’. In: *Nature Biotechnology* 42.2 (2024). Publisher: Nature Publishing Group, pp. 293–304.

[42] Mohammad Lotfollahi et al. ‘The future of rapid and automated single-cell data analysis using reference mapping’. English. In: *Cell* 187.10 (2024), pp. 2343–2358.

[43] Hans Clevers. ‘Modeling Development and Disease with Organoids’. In: *Cell* 165.7 (2016), pp. 1586–1597.

[44] Xuyu Qian et al. ‘Generation of human brain regionspecific organoids using a miniaturized spinning bioreactor’. In: *Nature Protocols* 3 (Mar. 2018), pp. 565–580.

[45] Anai Gonzalez-Cordero et al. 'Recapitulation of Human Retinal Development from Human Pluripotent Stem Cells Generates Transplantable Populations of Cone Photoreceptors'. In: *Stem Cell Reports* 9.3 (2017), pp. 820–837.

[46] Sergiu P. Paca. 'The rise of three-dimensional human brain cultures'. In: *Nature* 553.7689 (2018), pp. 437–445.

[47] Jessica M. Vanslambrouck et al. 'Generation of proximal tubule-enhanced kidney organoids from human pluripotent stem cells'. In: *Nature Protocols* 18.11 (2023), pp. 3229–3252.

[48] Bauer L. LeSavage et al. 'Next-generation cancer organoids'. In: *Nature Materials* 21.2 (2022), pp. 143–159.

[49] Minoru Takasato et al. 'Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis'. In: *Nature* 526.7574 (2015), pp. 564–568.

[50] Kim Bak Jensen and Melissa Helen Little. 'Organoids are not organs: Sources of variation and misinformation in organoid biology'. In: *Stem Cell Reports* 18.6 (2023), pp. 1255–1270.

[51] Yi Zhou, Hongjun Song and Guo-li Ming. 'Genetics of human brain development'. In: *Nature Reviews Genetics* 25.1 (2024). Publisher: Nature Publishing Group, pp. 26–45.

[52] Koen J. T. Venken, Julie H. Simpson and Hugo J. Bellen. 'Genetic Manipulation of Genes and Cells in the Nervous System of the Fruit Fly'. In: *Neuron* 72.2 (2011), pp. 202–230.

[53] Divya Jayaraman, Byoung-Il Bae and Christopher A. Walsh. 'The Genetics of Primary Microcephaly'. In: *Annual Review of Genomics and Human Genetics* 19 (2018), pp. 177–200.

[54] Madeline A. Lancaster. 'Unraveling mechanisms of human brain evolution'. In: *Cell* 187.21 (2024), pp. 5838–5857.

[55] Madeline A. Lancaster and Juergen A. Knoblich. 'Generation of cerebral organoids from human pluripotent stem cells'. In: *Nature Protocols* 9.10 (2014), pp. 2329–2340.

[56] Anca M. Paca et al. 'Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture'. In: *Nature Methods* 12.7 (2015), pp. 671–678.

[57] Ana Uzquiano et al. 'Proper acquisition of cell class identity in organoids allows definition of fate specification programs of the human cerebral cortex'. In: *Cell* 185.20 (2022), 3770–3788.e27.

[58] Silvia Velasco et al. 'Individual brain organoids reproducibly form cell diversity of the human cerebral cortex'. In: *Nature* 570.7762 (2019), pp. 523–527.

[59] Zhisong He et al. 'An integrated transcriptomic cell atlas of human neural organoids'. In: *Nature* 635.8039 (2024), pp. 690–698.

[60] Jaydeep Sidhaye et al. 'Integrated transcriptome and proteome analysis reveals posttranscriptional regulation of ribosomal genes in human brain organoids'. In: *eLife* 12 (2023), e85135.

[61] Anahita Amiri et al. 'Transcriptome and epigenome landscape of human cortical development modeled in organoids'. In: *Science* 362.6420 (2018), eaat6720.

[62] Charles A. Herring et al. 'Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution'. In: *Cell* 185.23 (2022), 4428–4447.e28.

[63] Norbert B. Ghyselinck and Gregg Duester. 'Retinoic acid signaling pathways'. In: *Development (Cambridge, England)* 146.13 (2019), dev167502.

[64] Gregg Duester. 'Towards a Better Vision of Retinoic Acid Signaling during Eye Development'. In: *Cells* 3 (2022), p. 322.

[65] C L Cepko et al. 'Cell fate determination in the vertebrate retina.' In: *Proceedings of the National Academy of Sciences* 2 (1996), pp. 589–595. (Visited on 15/09/2024).

[66] Michelle O'Hara-Wright and Anai Gonzalez-Cordero. 'Retinal organoids: a window into human retinal development'. In: *Development* 24 (2020), dev189746.

[67] E. M. Wells-Gray et al. 'Variation in rod and cone density from the fovea to the mid-periphery in healthy human retinas using adaptive optics scanning laser ophthalmoscopy'. In: *Eye* 30.8 (2016), pp. 1135–1143.

[68] C. A. Curcio et al. 'Human photoreceptor topography'. In: *The Journal of Comparative Neurology* 292.4 (1990), pp. 497–523.

[69] Susana da Silva and Constance L. Cepko. 'Fgf8 Expression and Degradation of Retinoic Acid Are Required for Patterning a High-Acuity Area in the Retina'. In: *Developmental Cell* 42.1 (2017), 68–81.e6.

[70] Xiufeng Zhong et al. 'Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs'. In: *Nature Communications* 1 (2014), p. 4047.

[71] Kiara C. Eldred et al. 'Thyroid hormone signaling specifies cone subtypes in human retinal organoids'. In: *Science* 362.6411 (2018), eaau6348.

[72] Sarah E. Hadyniak et al. 'Retinoic acid signaling regulates spatiotemporal specification of human green and red cones'. In: *PLOS Biology* 22.1 (2024), e3002464.

[73] George.A. Hyatt et al. 'Retinoic acid alters photoreceptor developmentinvivo'. In: *Proceedings of the National Academy of Sciences* 93.23 (1996), pp. 13298–13303.

[74] Ryoji Amamoto, Grace K. Wallick and Constance L. Cepko. 'Retinoic acid signaling mediates peripheral cone photoreceptor survival in a mouse model of retina degeneration'. In: *eLife* 11 (2022), e76389.

[75] Shun H. Yip, Pak Chung Sham and Junwen Wang. 'Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data'. In: *Briefings in Bioinformatics* 20.4 (2019), pp. 1583–1589.

[76] Aleksandra A. Kolodziejczyk et al. 'The Technology and Biology of Single-Cell RNA Sequencing'. In: *Molecular Cell* 58.4 (2015), pp. 610–620.

[77] Valentine Svensson, Sarah A Teichmann and Oliver Stegle. 'SpatialDE: identification of spatially variable genes'. In: *Nature Methods* 15.5 (2018), pp. 343–346.

[78] Shiquan Sun, Jiaqiang Zhu and Xiang Zhou. 'Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies'. In: *Nature Methods* 17.2 (2020), pp. 193–200.

[79] Jiaqiang Zhu, Shiquan Sun and Xiang Zhou. 'SPARK-X: non-parametric modeling enables scalable and robust detection of spatial expression patterns for large spatial transcriptomic studies'. In: *Genome Biology* 22.1 (2021), p. 184.

[80] Minsheng Hao, Kui Hua and Xuegong Zhang. 'SOMDE: A scalable method for identifying spatially variable genes with self-organizing map'. In: *Bioinformatics* (2021), btab471.

[81] Ruben Dries et al. 'Giotto: a toolbox for integrative analysis and visualization of spatial expression data'. In: *Genome Biology* 22.1 (2021), p. 78.

[82] Lukas M. Weber et al. 'nnSVG for the scalable identification of spatially variable genes using nearest-neighbor Gaussian processes'. In: *Nature Communications* 14.1 (2023), p. 4059.

[83] Brendan F. Miller et al. 'Characterizing spatial gene expression heterogeneity in spatially resolved single-cell transcriptomic data with nonuniform cellular densities'. In: *Genome Research* 31.10 (2021), pp. 1843–1855.

[84] Yuhan Hao et al. 'Integrated analysis of multimodal single-cell data'. In: *Cell* 184.13 (2021), 3573–3587.e29.

[85] Andrew L. Ji et al. 'Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma'. In: *Cell* 182.2 (2020), 497–514.e22.

[86] Samuel G. Rodriques et al. 'Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution'. In: *Science* 363.6434 (2019), pp. 1463–1467.

[87] Jamie L. Marshall et al. 'High-resolution Slide-seqV2 spatial transcriptomics enables discovery of disease-specific cell neighborhoods and pathways'. In: *iScience* 25.4 (2022).

[88] Chenglong Xia et al. 'Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression'. In: *Proceedings of the National Academy of Sciences* 116.39 (2019), pp. 19490–19499.

[89] S. Vickovic et al. 'SM-Omics is an automated platform for high-throughput spatial multi-omics'. In: *Nature Communications* 13.1 (2022), p. 795.

[90] Yang Liu et al. 'High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue'. In: *Cell* 183.6 (2020), 1665–1681.e18.

[91] P. A. P. Moran. 'NOTES ON CONTINUOUS STOCHASTIC PHENOMENA'. In: *Biometrika* 37.1-2 (1950), pp. 17–23.

[92] John L. Gittleman and Mark Kot. 'Adaptation: Statistics and a Null Model for Estimating Phylogenetic Effects'. In: *Systematic Zoology* 39.3 (1990), p. 227.

[93] Arkajyoti Saha and Abhirup Datta. 'BRISC: bootstrap for rapid inference on spatial covariances'. In: *Stat* 7.1 (2018), e184.

[94] Dongyuan Song et al. ‘scDesign3 generates realistic in silico data for multimodal single-cell and spatial omics’. In: *Nature Biotechnology* 42.2 (2024), pp. 247–252.

[95] Christian Thiele and Gerrit Hirschfeld. ‘**cutpointr** : Improved Estimation and Validation of Optimal Cutpoints in *R*’. In: *Journal of Statistical Software* 98.11 (2021).

[96] Edward Zhao et al. ‘Spatial transcriptomics at subspot resolution with BayesSpace’. In: *Nature Biotechnology* 39.11 (2021), pp. 1375–1384.

[97] Jian Hu et al. ‘SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network’. In: *Nature Methods* 18.11 (2021), pp. 1342–1351.

[98] Rui Jiang et al. ‘SINFONIA: Scalable Identification of Spatially Variable Genes for Deciphering Spatial Domains’. In: *Cells* 12.4 (2023), p. 604.

[99] Simone Romano et al. ‘Adjusting for Chance Clustering Comparison Measures’. In: *Journal of Machine Learning Research* 17.134 (2016), pp. 1–32.

[100] Chen Shengquan et al. ‘stPlus: a reference-based method for the accurate enhancement of spatial transcriptomics’. In: *Bioinformatics* 37.Supplement$_1$ (2021), pp. i299–i307.

[101] Yang Liu et al. *High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue*. 2020. URL: `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE137986`.

[102] Samuel G. Rodriques et al. *Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution*. 2019. URL: `https://singlecell.broadinstitute.org/single_cell/study/SCP354/slide-seq-study`.

[103] Robert R. Stickels et al. *Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2*. 2021. URL: `https://singlecell.broadinstitute.org/single_cell/study/SCP815/highly-sensitive-spatial-transcriptomics-at-near-cellular-resolution-with-slide-seqv2`.

[104] S. Vickovic et al. *SM-Omics is an automated platform for high-throughput spatial multi-omics*. 2022. URL: `https://singlecell.broadinstitute.org/single_cell/study/SCP979/sm-omics-an-automated-platform-for-high-throughput-spatial-multi-omics`.

[105] Andrew L. Ji et al. *P9_ST_rep1*. 2020. URL: `https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM4284322`.

[106] José Fernández Navarro et al. ‘Spatial Transcriptomics Reveals Genes Associated with Dysregulated Mitochondrial Functions and Stress Signaling in Alzheimer Disease’. In: *iScience* 23.10 (2020), p. 101556.

[107] Tommaso Biancalani et al. ‘Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram’. In: *Nature Methods* 18.11 (2021), pp. 1352–1362.

[108] Ricardo Melo Ferreira et al. ‘Integration of spatial and single-cell transcriptomics localizes epithelial cellimmune cross-talk in kidney injury’. In: *JCI Insight* 6.12 (2021), e147703.

[109] Miranda V. Hunter et al. ‘Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface’. In: *Nature Communications* 12.1 (2021), p. 6278.

[110] Danielle Janosevic et al. ‘The orchestrated cellular and molecular responses of the kidney to endotoxin define a precise sepsis timeline’. In: *eLife* 10 (2021), e62270.

[111] Anoushka Joglekar et al. ‘A spatially resolved brain region- and cell type-specific isoform atlas of the postnatal mouse brain’. In: *Nature Communications* 12.1 (2021), p. 463.

[112] Romain Lopez et al. ‘DestVI identifies continuums of cell types in spatial transcriptomics data’. In: *Nature Biotechnology* 40.9 (2022), pp. 1360–1369.

[113] Tara McCray et al. ‘Vitamin D sufficiency enhances differentiation of patient-derived prostate epithelial organoids’. In: *iScience* 24.1 (2021), p. 101974.

[114] Sunny Z. Wu et al. ‘A single-cell and spatially resolved atlas of human breast cancers’. In: *Nature Genetics* 53.9 (2021), pp. 1334–1347.

[115] Pengyi Yang, Hao Huang and Chunlei Liu. ‘Feature selection revisited in the single-cell era’. In: *Genome Biology* 22.1 (2021), p. 321.

[116] Shila Ghazanfar et al. ‘Investigating higher-order interactions in single-cell data with scHOT’. In: *Nature Methods* 17.8 (2020), pp. 799–806.

[117] Suoqin Jin et al. ‘Inference and analysis of cell-cell communication using CellChat’. In: *Nature Communications* 12.1 (2021), p. 1088.

[118] Mirjana Efremova et al. ‘CellPhoneDB: inferring cellcell communication from combined expression of multi-subunit ligandreceptor complexes’. In: *Nature Protocols* 15.4 (2020), pp. 1484–1506.

[119] Peter Langfelder and Steve Horvath. ‘WGCNA: an R package for weighted correlation network analysis’. In: *BMC Bioinformatics* 9 (2008), p. 559.

[120] Jonghun Kim, Gareth J. Sullivan and In-Hyun Park. ‘How well do brain organoids capture your brain?’ In: *iScience* 24.2 (2021), p. 102063.

[121] Kazutoshi Takahashi et al. ‘Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors’. In: *Cell* 131.5 (2007), pp. 861–872.

[122] Mototsugu Eiraku et al. ‘Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals’. In: *Cell Stem Cell* 3.5 (2008), pp. 519–532.

[123] Milan Fernando et al. ‘Differentiation of brain and retinal organoids from confluent cultures of pluripotent stem cells connected by nerve-like axonal projections of optic origin’. In: *Stem Cell Reports* 17.6 (2022), pp. 1476–1492.

[124] Ilaria Chiaradia and Madeline A. Lancaster. ‘Brain organoids for the study of human neurobiology at the interface of in vitro and in vivo’. In: *Nature Neuroscience* 23.12 (2020), pp. 1496–1508.

[125] Madeline A. Lancaster et al. ‘Cerebral organoids model human brain development and microcephaly’. In: *Nature* 501.7467 (2013), pp. 373–379.

[126] Silvia Velasco, Bruna Paulsen and Paola Arlotta. ‘3D Brain Organoids: Studying Brain Development and Disease Outside the Embryo’. In: *Annual Review of Neuroscience* 43.1 (2020), pp. 375–389.

[127] Joan Stiles and Terry L. Jernigan. ‘The Basics of Brain Development’. In: *Neuropsychology Review* 20.4 (2010), pp. 327–348.

[128] André M.M. Sousa et al. ‘Evolution of the Human Nervous System Function, Structure, and Development’. In: *Cell* 170.2 (2017), pp. 226–247.

[129] Matthew B. Johnson et al. ‘Functional and Evolutionary Insights into Human Brain Development through Global Transcriptome Analysis’. In: *Neuron* 62.4 (2009), pp. 494–509.

[130] Carlo Colantuoni et al. ‘Temporal dynamics and genetic control of transcription in the human prefrontal cortex’. In: *Nature* 478.7370 (2011), pp. 519–523.

[131] Hyo Jung Kang et al. ‘Spatio-temporal transcriptome of the human brain’. In: *Nature* 478.7370 (2011), pp. 483–489.

[132] Yoshiaki Tanaka et al. ‘Synthetic Analyses of Single-Cell Transcriptomes from Multiple Brain Organoids and Fetal Brain’. In: *Cell Reports* 30.6 (2020), 1682–1689.e3.

[133] Bradley J. Smith and Victor Corasolla Carregari. ‘Post-Translational Modifications During Brain Development’. In: *Advances in Experimental Medicine and Biology* 1382 (2022), pp. 29–38.

[134] Alexander Dobin et al. ‘STAR: ultrafast universal RNA-seq aligner’. In: *Bioinformatics* 29.1 (2013), pp. 15–21.

[135] Yang Liao, Gordon K Smyth and Wei Shi. ‘The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads’. In: *Nucleic Acids Research* 47.8 (2019), e47–e47.

[136] Yuqing Zhang, Giovanni Parmigiani and W Evan Johnson. ‘ComBat-seq: batch effect adjustment for RNA-seq count data’. In: *NAR Genomics and Bioinformatics* 2.3 (2020).

[137] Michael I Love, Wolfgang Huber and Simon Anders. ‘Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2’. In: *Genome Biology* 15.12 (2014), p. 550.

[138] Grace X. Y. Zheng et al. ‘Massively parallel digital transcriptional profiling of single cells’. In: *Nature Communications* 8.1 (2017), p. 14049.

[139] Jonathan A. Griffiths et al. ‘Detection and removal of barcode swapping in single-cell RNA-seq data’. In: *Nature Communications* 9.1 (2018), p. 2667.

[140] Christopher S. McGinnis, Lyndsay M. Murrow and Zev J. Gartner. ‘DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors’. In: *Cell Systems* 8.4 (2019), 329–337.e4.

[141] Kasper Engholm-Keller and Martin R. Larsen. ‘Improving the Phosphoproteome Coverage for Limited Sample Amounts Using TiO2-SIMAC-HILIC (TiSH) Phosphopeptide Enrichment and Fractionation’. In: *Methods in Molecular Biology (Clifton, N.J.)* 1355 (2016), pp. 161–177.

[142] Johannes Alexander Müller et al. ‘A presynaptic phosphosignaling hub for lasting homeostatic plasticity’. In: *Cell Reports* 39.3 (2022).

[143] Hani Jieun Kim et al. ‘PhosR enables processing and functional analysis of phosphoproteomic data’. In: *Cell Reports* 34.8 (2021), p. 108771.

[144] Pengyi Yang et al. ‘KinasePA: Phosphoproteomics data annotation using hypothesis driven kinase perturbation analysis’. In: *PROTEOMICS* 16.13 (2016), pp. 1868–1871.

[145] Pengyi Yang et al. ‘Knowledge-Based Analysis for Detecting Key Signaling Events from Time-Series Phosphoproteomics Data’. In: *PLOS Computational Biology* 11.8 (2015). Ed. by Lilia M. Iakoucheva, e1004403.

[146] Matthew E. Ritchie et al. ‘limma powers differential expression analyses for RNA-sequencing and microarray studies’. In: *Nucleic Acids Research* 43.7 (2015), e47–e47.

[147] Wen-Kang Shen et al. ‘AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations’. In: *Nucleic Acids Research* 51.D1 (2023), pp. D39–D45.

[148] Daria Marakulina et al. ‘EpiFactors 2022: expansion and enhancement of a curated database of human epigenetic factors and complexes’. In: *Nucleic Acids Research* D1 (2023), pp. D564–D570.

[149] Yingxin Lin et al. ‘scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets’. In: *Proceedings of the National Academy of Sciences* 116.20 (2019), pp. 9775–9784.

[150] Yingxin Lin et al. ‘Evaluating stably expressed genes in single cells’. In: *GigaScience* 8.9 (2019), giz106.

[151] Pengyi Yang et al. ‘Direction pathway analysis of large-scale proteomics data reveals novel features of the insulin action pathway’. In: *Bioinformatics* 30.6 (2014), pp. 808–814.

[152] Marc Gillespie et al. ‘The reactome pathway knowledgebase 2022’. In: *Nucleic Acids Research* 50.D1 (2022), pp. D687–D692.

[153] Raivo Kolde. *pheatmap: Pretty Heatmaps*. 2019. URL: `https://cran.r-project.org/web/packages/pheatmap/index.html`.

[154] Yichi Xu et al. ‘A single-cell transcriptome atlas profiles early organogenesis in human embryos’. In: *Nature Cell Biology* 25.4 (2023), pp. 604–615.

[155] Davis J. McCarthy et al. ‘Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R’. en. In: *Bioinformatics* (2017), btw777.

[156] Kristoffer T. G. Rigbolt et al. ‘System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation’. In: *Science Signaling* 4.164 (2011), rs3.

[157] I. Neganova et al. ‘CDK1 plays an important role in the maintenance of pluripotency and genomic stability in human pluripotent stem cells’. In: *Cell Death & Disease* 5.11 (2014), e1508–e1508.

[158] Matthew P. Swaffer et al. ‘CDK Substrate Phosphorylation and Ordering the Cell Cycle’. In: *Cell* 167.7 (2016), 1750–1761.e16.

[159] K. L. Guan et al. ‘Negative regulation of the serine/threonine kinase B-Raf by Akt’. In: *The Journal of Biological Chemistry* 275.35 (2000), pp. 27354–27359.

[160] Mitchell Cheung et al. ‘Akt3 and mutant V600E B-Raf cooperate to promote early melanoma development’. In: *Cancer Research* 68.9 (2008), pp. 3429–3439.

[161] Martin Schmidlin et al. ‘The ARE-dependent mRNA-destabilizing activity of BRF1 is regulated by protein kinase B’. In: *The EMBO Journal* 23.24 (2004), pp. 4760–4769.

[162] Chao Feng et al. ‘TCF20 dysfunction leads to cortical neurogenesis defects and autisticlike behaviors in mice’. In: *EMBO reports* 21.8 (2020).

[163] Elif Nur Firat-Karalar, Peter P. Hsiue and Matthew D. Welch. ‘The actin nucleation factor JMY is a negative regulator of neuritogenesis’. In: *Molecular Biology of the Cell* 22.23 (2011), pp. 4563–4574.

[164] Guang Jin Pan et al. ‘Stem cell pluripotency and transcription factor Oct4’. In: *Cell Research* 12.5-6 (2002), pp. 321–329.

[165] Xin Zhang et al. ‘FOXO1 is an essential regulator of pluripotency in human embryonic stem cells’. In: *Nature Cell Biology* 13.9 (2011), pp. 1092–1099.

[166] Masaki Okano et al. ‘DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development’. In: *Cell* 99.3 (1999), pp. 247–257.

[167] D. Watanabe, K. Uchiyama and K. Hanaoka. ‘Transition of mouse de novo methyltransferases expression from Dnmt3b to Dnmt3a during neural progenitor cell development’. In: *Neuroscience* 142.3 (2006), pp. 727–737.

[168] Chris T. Dee et al. ‘Sox3 regulates both neural fate and differentiation in the zebrafish ectoderm’. In: *Developmental Biology* 320.1 (2008), pp. 289–301.

[169] Yechiel Elkabetz et al. ‘Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage’. In: *Genes & Development* 22.2 (2008), pp. 152–165.

[170] Xiaoqing Zhang et al. ‘Pax6 Is a Human Neuroectoderm Cell Fate Determinant’. In: *Cell Stem Cell* 7.1 (2010), pp. 90–100.

[171] Deepak Kumar, Erez Nitzan and Chaya Kalcheim. ‘YAP promotes neural crest emigration through interactions with BMP and Wnt activities’. In: *Cell Communication and Signaling* 17.1 (2019), p. 69.

[172] Yi-Ting Lin et al. ‘YAP regulates neuronal differentiation through Sonic hedgehog signaling pathway’. In: *Experimental Cell Research* 318.15 (2012), pp. 1877–1888.

[173] Hui Wang et al. ‘Gli3 Is Required for Maintenance and Fate Specification of Cortical Progenitors’. In: *The Journal of Neuroscience* 31.17 (2011), pp. 6440–6448.

[174] Zachary B. Gaber, Samantha J. Butler and Bennett G. Novitch. ‘PLZF Regulates Fibroblast Growth Factor Responsiveness and Maintenance of Neural Progenitors’. In: *PLoS Biology* 11.10 (2013), e1001676.

[175] Eitan Shaulian and Michael Karin. ‘AP-1 as a regulator of cell life and death’. In: *Nature Cell Biology* 4.5 (2002), E131–E136.

[176] Begoña Canovas and Angel R. Nebreda. ‘Diversity and versatility of p38 kinase signalling in health and disease’. In: *Nature Reviews Molecular Cell Biology* 22.5 (2021), pp. 346–366.

[177] Ruben K. Dagda and Tania Das Banerjee. ‘Role of PKA in regulating mitochondrial function and neuronal development: implications to neurodegenerative diseases’. In: *Reviews in the neurosciences* 26.3 (2015), pp. 359–370.

[178] Josien Levenga et al. ‘Immunohistological Examination of AKT Isoforms in the Brain: Cell-Type Specificity That May Underlie AKTs Role in Complex Brain Disorders and Neurological Disease’. In: *Cerebral Cortex Communications* 2.2 (2021), tgab036.

[179] Sigrid Uxa et al. ‘Ki-67 gene expression’. In: *Cell Death & Differentiation* 28.12 (2021), pp. 3357–3370.

[180] Clifford D.L. Folmes et al. ‘Metabolic Plasticity in Stem Cell Homeostasis and Differentiation’. In: *Cell Stem Cell* 11.5 (2012), pp. 596–606.

[181] Hakryul Jo et al. ‘Small molecule-induced cytosolic activation of protein kinase Akt rescues ischemia-elicited neuronal death’. In: *Proceedings of the National Academy of Sciences* 109.26 (2012), pp. 10581–10586.

[182] Harvey B. Sarnat. ‘Immunocytochemical markers of neuronal maturation in human diagnostic neuropathology’. In: *Cell and Tissue Research* 359.1 (2015), pp. 279–294.

[183] Noriko Osumi et al. ‘Concise Review: Pax6 Transcription Factor Contributes to both Embryonic and Adult Neurogenesis as a Multifunctional Regulator’. In: *STEM CELLS* 26.7 (2008), pp. 1663–1672.

[184] William L. Hatleberg and Veronica F. Hinman. ‘Modularity and hierarchy in biological systems: Using gene regulatory networks to understand evolutionary change’. In: *Current Topics in Developmental Biology* 141 (2021), pp. 39–73.

[185] A. J. Whitmarsh and R. J. Davis. ‘Regulation of transcription factor function by phosphorylation’. In: *Cellular and molecular life sciences: CMLS* 57.8-9 (2000), pp. 1172–1183.

[186] Josien Levenga et al. ‘AKT isoforms have distinct hippocampal expression and roles in synaptic plasticity’. In: *eLife* 6 (2017), e30640.

[187] Runrui Zhang, Anna Engler and Verdon Taylor. ‘Notch: an interactive player in neurogenesis and disease’. In: *Cell and Tissue Research* 371.1 (2018), pp. 73–89.

[188] Yun Li et al. ‘Induction of Expansion and Folding in Human Cerebral Organoids’. In: *Cell Stem Cell* 20.3 (2017), 385–396.e3.

[189] Richard H. Masland. ‘The neuronal organization of the retina’. eng. In: *Neuron* 76.2 (2012), pp. 266–280.

[190] Gregg Duester. ‘Retinoic Acid Synthesis and Signaling during Early Organogenesis’. In: *Cell* 6 (2008), pp. 921–931.

[191] M. W. Kelley, J. K. Turner and T. A. Reh. ‘Retinoic acid promotes differentiation of photoreceptors in vitro’. In: *Development (Cambridge, England)* 120.8 (1994), pp. 2091–2102.

[192] Katarzyna A. Hussey, Sarah E. Hadyniak and Robert J. Johnston. ‘Patterning and Development of Photoreceptors in the Human Retina’. In: *Frontiers in Cell and Developmental Biology* 10 (2022).

[193] Hani Jieun Kim et al. ‘Comprehensive characterization of fetal and mature retinal cell identity to assess the fidelity of retinal organoids’. In: *Stem Cell Reports* 18.1 (2023), pp. 175–189.

[194] Akshayalakshmi Sridhar et al. ‘Single-Cell Transcriptomic Comparison of Human Fetal Retina, hPSC-Derived Retinal Organoids, and Long-Term Retinal Cultures’. In: *Cell Reports* 30.5 (2020), 1644–1659.e4.

[195] Alyssa Kallman et al. ‘Investigating cone photoreceptor development using patient-derived NRL null retinal organoids’. In: *Communications Biology* 3.1 (2020), pp. 1–13.

[196] Cameron S. Cowan et al. ‘Cell Types of the Human Retina and Its Organoids at Single-Cell Resolution’. In: *Cell* 182.6 (2020), 1623–1640.e34.

[197] Birthe Dorgau et al. ‘Deciphering the spatiotemporal transcriptional and chromatin accessibility of human retinal organoid development at the single-cell level’. In: *iScience* 27.4 (2024).

[198] Birthe Dorgau et al. ‘Single-cell analyses reveal transient retinal progenitor cells in the ciliary margin of developing human retina’. In: *Nature Communications* 15.1 (2024), p. 3567.

[199] Philipp Wahle et al. ‘Multimodal spatiotemporal phenotyping of human retinal organoid development’. In: *Nature Biotechnology* 41.12 (2023), pp. 1765–1775.

[200] Valentin M. Sluch et al. ‘Differentiation of human ESCs to retinal ganglion cells using a CRISPR engineered reporter cell line’. In: *Scientific Reports* 5.1 (2015), p. 16595.

[201] Emma L. West et al. ‘Antioxidant and lipid supplementation improve the development of photoreceptor outer segments in pluripotent stem cell-derived retinal organoids’. In: *Stem Cell Reports* 17.4 (2022), pp. 775–788.

[202] Zhen Zuo et al. ‘Single cell dual-omic atlas of the human developing retina’. In: *Nature Communications* 15.1 (2024), p. 6792.

[203] Vinay S Swamy et al. ‘Building the mega single-cell transcriptome ocular meta-atlas’. In: *GigaScience* 10.10 (2021), giab061.

[204] Kim Hj et al. ‘Uncovering cell identity through differential stability with Cepo’. In: *Nature computational science* 1.12 (2021).

[205] Craig B. Stevens, David A. Cameron and Deborah L. Stenkamp. ‘Plasticity of photoreceptor-generating retinal progenitors revealed by prolonged retinoic acid exposure’. In: *BMC Developmental Biology* 11.1 (2011), p. 51.

[206] Albane le Maire et al. ‘A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor’. In: *Nature Structural & Molecular Biology* 17.7 (2010), pp. 801–807.

[207] Yuval Rinkevich et al. ‘Systemic Bud Induction and Retinoic Acid Signaling Underlie Whole Body Regeneration in the Urochordate Botrylloides leachi’. In: *PLOS Biology* 5.4 (2007), e71.

[208] Yufeng Lu et al. ‘Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development’. In: *Developmental Cell* 53.4 (2020), 473–491.e9.

[209] Wenyang Yi et al. ‘A single-cell transcriptome atlas of the aging human and macaque retina’. In: *National Science Review* 8.4 (2021), nwaa179.

[210] Jade Harkin et al. ‘A highly reproducible and efficient method for retinal organoid differentiation from human pluripotent stem cells’. In: *Proceedings of the National Academy of Sciences* 121.25 (2024), e2317285121.

[211] Kazutoshi Takahashi and Shinya Yamanaka. ‘Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors’. In: *Cell* 126.4 (2006), pp. 663–676.

[212] Gaoyang Liang and Yi Zhang. ‘Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective’. In: *Cell Research* 23.1 (2013), pp. 49–69.

[213] Xuyu Qian, Hongjun Song and Guo-li Ming. ‘Brain organoids: advances, applications and challenges’. In: *Development* 146.8 (2019), dev166074.

[214] Yue You et al. ‘Systematic comparison of sequencing-based spatial transcriptomic methods’. In: *Nature Methods* 21.9 (2024), pp. 1743–1754.

[215] Jonas Simon Fleck et al. ‘Inferring and perturbing cell fate regulomes in human brain organoids’. In: *Nature* 7978 (2023), pp. 365–372.

[216] Jiho Choi et al. ‘Characterization of the development of the high-acuity area of the chick retina’. In: *Developmental Biology* 511 (2024), pp. 39–52.

[217] Madison R. Glass et al. ‘Cross-site reproducibility of human cortical organoids reveals consistent cell type composition and architecture’. In: *Stem Cell Reports* 19.9 (2024), pp. 1351–1367.