Epigenetics & Chromatin

RESEARCH Open Access

Chromatin profiling and state predictions reveal insights into epigenetic regulation during early porcine development

Sarah M. Innis[1] and Ryan A. Cabot[1*]

Abstract

Background Given their physiological similarities to humans, pigs are increasingly used as model organisms in human-oriented biomedical studies. Additionally, their value to animal agriculture across the globe has led to the development of numerous studies to investigate how to improve livestock welfare and production efficiency. As such, pigs are uniquely poised as compelling models that can yield findings with potential implications in both human and animal contexts. Despite this, many gaps remain in our knowledge about the foundational mechanisms that govern gene expression in swine across different developmental stages, particularly in early development. To address some of these gaps, we profiled the histone marks H3K4me3, H3K27ac, and H3K27me3 and the SWI/SNF central ATPase BRG1 in two porcine cell lines representing discrete early developmental time points and used the resulting information to construct predicted chromatin state maps for these cells. We combined this approach with analysis of publicly available RNA-seq data to examine the relationship between epigenetic status and gene expression in these cell types.

Results In porcine fetal fibroblast (PFF) and trophectoderm cells (PTr2), we saw expected patterns of enrichment for each of the profiled epigenetic features relative to specific genomic regions. H3K4me3 was primarily enriched at and around global gene promoters, H3K27ac was enriched in promoter and intergenic regions, H3K27me3 had broad stretches of enrichment across the genome and narrower enrichment patterns in and around the promoter regions of some genes, and BRG1 primarily had detectable enrichment at and around promoter regions and in intergenic stretches, with many instances of H3K27ac co-enrichment. We used this information to perform genome-wide chromatin state predictions for 10 different states using ChromHMM. Using the predicted chromatin state maps, we identified a subset of genomic regions marked by broad H3K4me3 enrichment, and annotation of these regions revealed that they were highly associated with essential developmental processes and consisted largely of expressed genes. We then compared the identities of the genes marked by these regions to genes identified as cell-type-specific using transcriptome data and saw that a subset of broad H3K4me3-marked genes was also specifically expressed in either PFF or PTr2 cells.

Conclusions These findings enhance our understanding of the epigenetic landscape present in early swine development and provide insight into how variabilities in chromatin state are linked to cell identity. Furthermore, this

*Correspondence:
Ryan A. Cabot
rcabot@purdue.edu

Full list of author information is available at the end of the article

data captures foundational epigenetic details in two valuable porcine cell lines and contributes to the growing body of knowledge surrounding the epigenetic landscape in this species.

Keywords Epigenetics, Histone modifications, Chromatin remodeling, CUT&RUN, Development

Background

Epigenetics involves the study of the genome-wide presence of heritable gene expression modifiers capable of influencing transcription without changing the existing DNA sequence. Histone modifications such as methylation and acetylation represent a significant group of epigenetic features within the epigenome, and together with chromatin remodeling complexes, they contribute to the precise control of chromatin architecture and, by extension, the accessibility of genes to transcriptional machinery. Studying how the epigenome regulates gene expression during the earliest stages of organism development can reveal findings with broad implications for developmental biology, disease research, and regenerative medicine.

As a model organism, the pig is conveniently poised to contribute insights into both human health and animal agriculture. Pigs are an increasingly valuable biomedical model due to their similarities in mature organ size, metabolic physiology, and immune function to humans, especially as xenotransplantation of pig-derived tissues grows in prevalence of practice [1, 2]. In addition, several transgenic pig lines have been established to facilitate the study of several human diseases, including cystic fibrosis, diabetes mellitus, and various neurodegenerative and cardiovascular conditions [3–9], many of which were established via somatic cell nuclear transfer (SCNT) using cultured porcine fetal fibroblasts (PFFs) as nuclear donors. As a food animal, pigs remain a staple livestock species and food source worldwide, and recent research has endeavored to explore the creation of pigs that are faster growing, more feed-efficient, and less susceptible to stress and disease [10–12]. This includes studying the roles of the uterus and placenta in conferring nutrition *in utero* and providing an environment conducive to embryonic and fetal growth. Indeed, several studies focusing on increasing our understanding of placental physiology and gestational pathologies have been performed using the pig as a model organism, many of which have utilized a line of porcine trophectoderm (PTr2) cells for in vitro experiments [13–17].

Despite the promise of their utility as a biomedical model and widely-established value in food animal agriculture, many knowledge gaps remain concerning the porcine epigenome, particularly across different developmental time points. While a significant portion of porcine epigenome information has been obtained using information from DNA methylation studies, some efforts have been made to develop chromatin landscape maps across different swine tissues [18–21], especially as part of the broader Functional Annotation of Animal Genomes (FAANG) initiative [22]. However, to date, few studies have evaluated the epigenetic landscape or provided functional annotation details in porcine cells from early developmental time points. To address this dearth of information and contribute to growing efforts to generate epigenetic annotation data in swine models, we used CUT&RUN followed by next-generation sequencing (NGS) to evaluate the state of the histone modifications H3K4me3, H3K27ac, H3K27me3, and the evolutionarily-conserved SWI/SNF chromatin remodeling complex central ATPase BRG1 (also known as SMARCA4) in PTr2 and PFF cells. Using this data, we constructed genome-wide predicted chromatin state map for these cell lines. From these findings, we identified regions of the genome marked with broad H3K4me3 signal and annotated them to examine their potential roles in governing cell function and identity. Taken together, the present research provides an in vitro basis for obtaining epigenetic information in two discrete early developmental time points in swine and enhances our understanding of how specific histone modification patterns may guide cell identity and function.

Methods

Unless otherwise stated, all chemicals were procured from Sigma Chemical Company, St. Louis, MO, USA. All cell culture reagents were obtained from Thermo Fisher Scientific, Waltham, MA, USA. Furthermore, with the exception of some antibodies, all CUT&RUN reaction reagents were purchased from EpiCypher, Durham, NC, USA.

Cell lines and culture conditions

PFFs were harvested from a porcine conceptus on day 40 of gestation as previously described [23]. Cells were cultured in DMEM supplemented with 15% FBS, 1% L-glutamine, 1% sodium pyruvate, 1% MEM-nonessential amino acids, and 1% penicillin-streptomycin and then collected at passage 5 for use in this study. Non-primary PTr2 cells were grown as previously described [24] in phenol red-free DMEM-F12 supplemented with 5% FBS, 0.1 units/mL bovine insulin, 1% L-glutamine, and 1% penicillin-streptomycin. Both cell lines were cultured in 150 cm^2 cell culture flasks at 39°C, 5% CO_2, and 100% humidity. Cells were passaged at 90% confluency using trypsin/EDTA and were at passage 60 at the time of use in these experiments.

Antibodies

The following antibodies were used in this study: CUTANA Rabbit IgG (13–0042, EpiCypher, Durham, NC), Anti-Histone H3 trimethyl K4 (13–0041, EpiCypher), Anti-BRG1/SMARCA4 (13-2002, EpiCypher), Anti-Histone H3 acetyl K27 (ab4729, Abcam, Cambridge, United Kingdom), Anti-Histone H3 trimethyl K27 (07-449, Sigma).

CUT&RUN

The CUT&RUN assay was performed as previously described [25]. To elaborate, cells were harvested using trypsin/EDTA, counted on a hemacytometer, and then aliquoted at 5×10^5 cells per sample onto magnetic Concanavalin A beads (21-1401). For each sample of immobilized cells, 0.5 µg of primary antibody was added, and the samples were incubated overnight on an orbital rotator (Boekel Scientific, Philadelphia, PA, USA) at 4 °C. Following incubation, bead-bound cells were permeabilized, and 2.5 µL of protein A/G micrococcal nuclease (pAG-MNase) (15-1016) was added per sample and incubated for 10 min at room temperature (RT). The pAG-MNase was activated by the addition of 1 µL 100 nM $CaCl_2$ (21-1007), and samples were incubated on an orbital rotator for 2 h at 4 °C to allow the chromatin cleavage reaction to proceed. At the end of the 2-hour incubation period, pAG-MNase activity was terminated by the addition of Stop Buffer (21-1003) containing *E. coli* Spike-in DNA (18-1401) and a 10-minute incubation period at 37 °C. After this incubation period, CUT&RUN enriched DNA was purified using DNA Binding Buffer (21-1008), DNA Wash Buffer (21-1009), and DNA Elution Buffer (21-1010) provided in the CUT&RUN kit. The purified CUT&RUN DNA was then quantified on a Qubit fluorometer (ThermoFisher) before proceeding to library preparation. Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (E7645S, New England Biolabs (NEB), Ipswich, MA, USA), NEBNext Multiplex Oligos for Illumina (Sets 1–4) (Set 1: E7335S, Set 2: E7500S, Set 3: E7710S, Set 4: E7730S, NEB), and Agencourt AMPure XP Beads (A63880, Beckman Coulter, Brea, CA). Libraries were quantified on a Qubit fluorometer and were evaluated for size on a TapeStation system (Agilent Technologies, Santa Clara, CA, USA). Paired-end sequencing (150 bp read length) was performed on an Illumina NovaSeq 6000 system (Illumina, San Diego, CA, USA) at a read depth of 5–8 million reads per library.

CUT&RUN sequencing data analysis

Adapter sequences were removed from demultiplexed, paired-end Illumina reads (as FASTQ files) using TrimGalore (v. 0.6.7) [26] with the - -fastqc parameter. Trimmed reads were aligned to the pig genome assembly Sscrofa11.1 (GenBank accession GCA-000003025.6) and the *E. coli* K12, MG1655 reference genome (obtained from https://support.illumina.com/sequencing/sequencing_software/igenome.html) using Bowtie2 (v. 2.4.5) [27] and the options --local, --very-sensitive, --no-unal, --no-mixed, and --no-discordant. The resulting SAM files (aligned to Sscrofa11.1) were converted to BAM files, sorted by genomic coordinates, and indexed using SAMtools (v. 1.15.1) [28] sort and index, respectively. BAM files were filtered to remove unmapped reads, improperly paired reads, and reads with a MAPQ score < 20 using SAMtools view. Using the *E. coli* alignment files, normalization factors were calculated as previously described [29]. Normalized bigWig files were created using deepTools (v. 3.5.1) [30] bamCoverage with the --scaleFactor option set to the corresponding normalization factor calculated previously, a --binSize of 20, and a --smoothLength of 60. Additionally, reads were extended and centered. Coverage tracks were visualized using the Broad Institute's Integrated Genomics Viewer (IGV) [31]. A BED file containing the gene coordinates for the Sscrofa11.1 assembly was obtained using the UCSC table browser [32]. Reference point matrices were generated using deepTools computeMatrix and profile plot and heatmap visualizations of these matrices were generated using deeptools plotProfile and plotHeatmap, respectively. MACS2 was used to call peaks with - - format BAMPE and - - keep-dup all, and corresponding IgG BAM files were included as controls [33]. The - - broad parameter was included for H3K27me3 samples. Peak detection was passed on a minimum FDR cutoff (q-value) of 0.05. Bedtools (v. 2.30.0) intersect was used to compare replicate and sample peaks. Replicate PCA plots and correlation heatmaps were created using the R Bioconductor package DiffBind (v. 3.8.4) using read count data normalized by library size [34, 35]. Binding site overlaps between replicates and samples were also evaluated using DiffBind. Binding site annotation was performed using the Bioconductor annotation packages ChIPseeker (v. 1.34.1) [36] and ChIPpeakAnno (v. 3.32.0) [37]. For Gene Ontology overrepresentation testing, one-sided Fisher's exact test was used to test for statistical significance ($p \leq 0.05$), and Benjamini–Hochberg correction (FDR) was performed to obtain adjusted p-values ($q \leq 0.05$). For chromatin state predictions, BAM files were converted to BED with Bedtools bamtobed and binarized using BinarizeBed from ChromHMM (v. 1.25) [38]. Models were generated using the ChromHMM LearnModel command, and output BED files were visualized in IGV. Motif enrichment analysis and determination of enriched motif p-values using cumulative binomial distribution were performed using findMotifsGenome.pl in HOMER with default parameters.

RNA-seq data analysis

RNA-seq data were obtained from the NIH BioProject database. Raw FASTQ files were sourced from accessions PRJNA798047 and PRJNA778857 and then aligned to the Sscrofa11.1 reference genome (NCBI RefSeq assembly GCF_000003025.6). Sscrofa11.1 was indexed with the hisat2-build command, and read alignments were performed using HISAT2 (v. 2.2.1) [39, 59]. Counts were obtained with HTSeq-count [40, 60] using the NCBI Sscrofa11.1 annotation features (GTF) genome file and with the mode option set to intersection-nonempty. Limma [41, 61] was used to obtain normalized read count data for each count file using a *P* value adjusted threshold of 0.05 (Benjamini–Hochberg correction) and trimmed mean of M values (TMM) normalization [42]. Tissue-specific genes were identified as previously described [43, 44]. Briefly, a biological category (i.e. fibroblast) was assigned to each sample according to material source for each species. A *t*-statistic was calculated for each gene by excluding expression data belonging to other cells in the same biological category, then each gene list was ranked by *t*-statistic. The top 10% of genes as ranked by *t*-statistic were defined as likely tissue-specific genes and used for downstream comparisons between cell types.

RT-qPCR characterization of PTr2 cells

RNA was extracted from PFF (passage 5) and PTr2 (passage 60) cells with the Direct-zol RNA Miniprep kit (Zymo Research, Irvine, CA, USA). Total RNA was quantified on a Nanodrop ND-1000 (ThermoFisher) device, and integrity was evaluated via denaturing gel. RNA was reverse-transcribed with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). Primers targeting genes of interest were designed with the NCBI Primer-BLAST (Table 1) and were produced by IDT (Integrated DNA Technologies, Coralville, IA, USA). All primer validation and qPCR runs were performed on a CFX Connect Real-Time System (Bio-Rad, Hercules, CA, USA). Primer efficiency for each target was between 95 and 105%, with no evidence of multiple amplicons seen in the melt curve analysis. For relative quantification of gene expression, qPCR was performed in technical triplicate for each cell type with the SsoAdvanced Universal SYBR Green SuperMix (Bio-Rad). Transcript data was normalized to GAPDH, and PFF cells were used as the calibrator. The $2^{-\Delta\Delta Ct}$ method was used to obtain relative log fold change expression values for PTr2 cells.

Results

Global enrichment patterns of H3K4me3, H3K27ac, H3K27me3, and BRG1 in PFF cells

H3K4me3 was strongly enriched around transcriptional start sites (TSSs) and had minimal detectable enrichment outside of these regions (Fig. 1A). The majority of H3K27ac signal was also present around global TSSs, but a larger proportion of H3K27ac binding was apparent outside TSS and gene bodies relative to H3K4me3 (Fig. 1B). Some H3K27me3 signal was found proximal to TSS regions, but enrichment was also broadly present outside of known gene coordinates (Fig. 1C). BRG1 signal showed a similar enrichment pattern to that of H3K27ac, albeit with lower overall signal abundance (Fig. 1D). Overall, these observations were supported by genomic region annotation of the selected epigenetic features, and the largest shares of each annotation category are included here. Nearly 90% (87.4%) of H3K4me3 signal was in gene promoter regions (Fig. 1E), consistent with the expected enrichment characteristics for this mark. Of the regions identified as represented in the H3K27ac annotation, the largest share of H3K27ac signal was present in promoter regions (41.58%), though just under a third (32.88%) of H3K27ac sites were in introns, and 12.40% were in distal intergenic regions (Fig. 1F). Approximately 22% of H3K27me3 signal was detected in promoter regions, while 46.33% of signal was in distal intergenic regions (Fig. 1G). For BRG1, 14.85% of signal was in promoter regions, while 41.28% was in introns, and around a third (33.43%) was in distal intergenic regions. Gene ontology (GO) analysis of each of these features showed that their enrichment was associated with a broad range of biological processes, including multiple terms concerned with key aspects of developmental progression (Fig. 1I-L). Furthermore, comparison

Table 1 Primer sequences for genes of interest used for qPCR validation of PTr2 cells

Target Gene	RefSeq Gene ID	Forward Primer	Reverse Primer	Annealing Temperature (°C)	Amplicon Size (bp)
FGFR2/KGFR	396762	5′-AAGATGATGCCACAGAGAAAGA-3′	5′-CAGGCTCCGAGGAGATTTATG-3′	62	103
GAPDH	110260753	5′-GGTGAAGGTCGGAGTGAACG-3′	5′-TGACTGTGCCGTGGAATTTG-3′	62	101
KRT7	100626722	5′-ACCAGACCAAGTTTGAGACC-3′	5′-TCGGTTCATCTCCGCAATC-3′	62	103
KRT8	100152077	5′-GGTTCTGGAGACCAAATGGAA-3′	5′-CGCCGGAGGTTGTTGATATAG-3′	62	96
ITGA4	100521477	5′-GGTGGCTGGAGAATGAGAAA-3′	5′-ACTGGTACACACCAAGTTAAGG-3′	62	103
ITGA5	100155091	5′-TCCTACATTACCAGAGCAAGAG-3′	5′-CAGGTCTGGCACACAGATATT-3′	62	90
ITGB1	397019	5′-ACCTTATGGACCTCTCCTACTC-3′	5′-ACTCTGAAGTAATCCTCCTCATTTC-3′	62	99

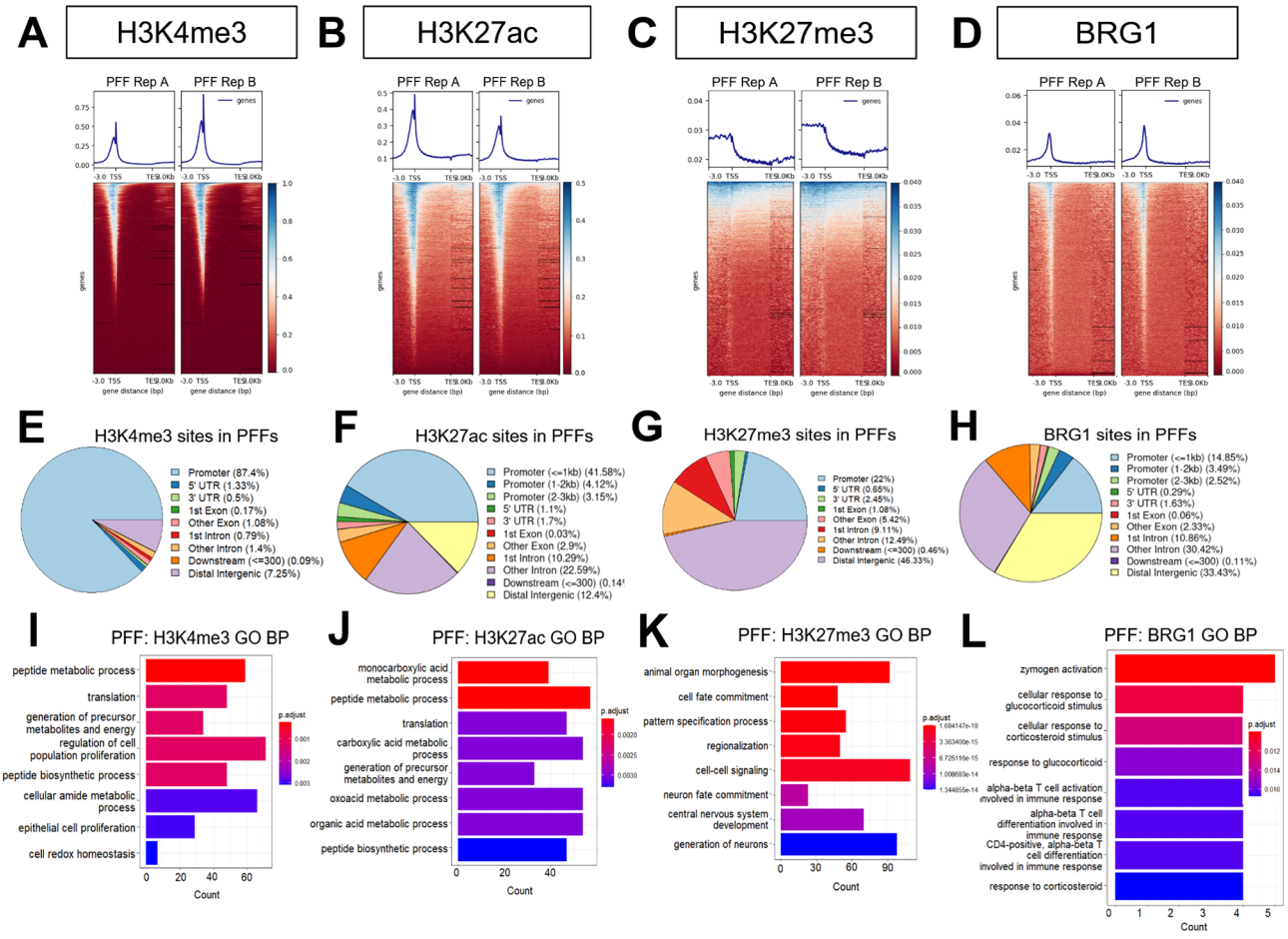

Fig. 1 Genomic localization and functional annotation of epigenetic features in PFF cells. Enrichment of (**A**) H3K4me3, (**B**) H3K27ac, (**C**) H3K27me3, and (**D**) BRG1 relative to global TSS and TES locations. Genomic feature annotation of global (**E**) H3K4me3, (**F**) H3K27ac, (**G**) H3K27me3, and (**H**) BRG1 binding. GO biological process overrepresentation analysis ($P < 0.05$ using one-sided Fisher's exact test; adjusted $P < 0.05$ after Benjamin-Hochberg FDR) for (**I**) H3K4me3, (**J**) H3K27ac, (**K**) H3K27me3, and (**L**) BRG1

of the epigenetic data to publicly-available RNA-seq data showed that the genes identified in the top 3 GO biological processes for H3K4me3 and H3K27ac were largely expressed and had higher expression levels relative to genes enriched for H3K27me3 (Supplementary Fig. S1A-C).

Characterization of PTr2 cells

As the PTr2 cells used in this work were at a later passage at the time of CUT&RUN, we first wanted to characterize these cells using select marker genes. Both the cytokeratin genes *KRT7* and *KRT8* were included in this characterization, as they have both been shown to be present in PTr2 cells [45, 46]. Additionally, the gene keratinocyte growth factor receptor (*KGFR*) gene, also known as FGFR2, for fibroblast growth factor receptor 2 (*FGFR2)* was included, as this gene is known to be expressed in PTr2s [46, 47]. Several integrins have also been shown to be expressed in PTr2 cells, such as *ITGA5* [48, 49] and *ITGB1* [49]. It has been reported, however, that the integrin *ITGA4* is not expressed in this cell line [49], so this gene target was also included to further validate the character of these PTr2 cells. Exploration of publicly-available RNA-seq data for this cell line confirmed that *KRT7* and *KRT8* were both highly expressed in this cell line, as was *ITGB1* (Supplementary Fig. S2A). *FGFR2* also appeared to be expressed, albeit at a lower level than the keratin genes. *ITGA5* expression data suggested that this gene is likely lowly expressed in PTr2s, while *ITGA4* is not expressed (Supplementary Fig. S2A). *RT*-qPCR data confirmed that both *KRT7* and *KRT8* were highly expressed in the PTr2 cells, consistent with what would be expected for trophectoderm cells given their epithelial character (Supplementary Fig. S2B). Analysis of transcriptional data for *KGFR/FGFR2* indicated that this gene was more moderately expressed in PTr2 cells than *KRT7* and *KRT8* and that expression levels were similar between PTr2 and PFF cells. For the integrins, *ITGA5* appeared to have a considerably lower level of expression in PTr2 cells relative to PFFs. Previous research has

shown that while *ITGA5* expression is detectable in PTr2 cells, it is comparatively lower than that of other integrins such as *ITGB1* [49]. Indeed, while PTr2 and PFF cells had a similar level of *ITGB1* expression, examination of the amplification data for this gene indicated that it was more highly expressed than *ITGA5* in PTr2s (Supplementary Fig. S2B). *ITGA4* expression was very low in PTr2 cells, and this observation is consistent with what has been reported in the literature [49]. Taken together, these results indicate that the PTr2 cells used in this study have retained PTr2 character despite their passage status.

Global enrichment patterns of H3K4me3, H3K27ac, H3K27me3, and BRG1 in PTr2 cells

Enrichment patterns for the selected epigenetic features were broadly similar between PFF and PTr2 cells. H3K4me3 signal was again strongly enriched around global TSS locations (Fig. 2A). H3K27ac signal was clearly detected at and immediately around TSSs, and the presence of this mark outside of gene coordinates was also apparent (Fig. 2B). As with PFF cells, H3K27me3 signal showed a small spike in enrichment around TSSs, but overall, signal for this mark seemed to be primarily outside of TSS and gene bodies (Fig. 2C). Additionally, PTr2 BRG1 signal also showed clear enrichment around TSSs, though heatmap patterns suggested that a majority of BRG1 signal was likely present outside of these sites (Fig. 2D). Genomic region annotation of H3K4me3 showed comparable results to what was seen in PFFs cells, with 82.72% of enrichment sites for this mark detected in promoter regions. A larger share of H3K27ac signal was detected in PTr2 cell promoters (57.36%) than what was observed in PFF cells, though less disparity was seen in the proportion of signal split between introns and distal intergenic regions PTr2s, with 20.77% and 12.22% of H3K27ac signal detected in these regions, respectively (Fig. 2F). H3K27me3 signal in PTr2 promoter regions was lower than that of PFFs, at 7.77%, while the proportion of signal detected in distal intergenic regions was highly similar, at 48.3% (Fig. 2G). For BRG1 in PTr2 cells, 21.49%

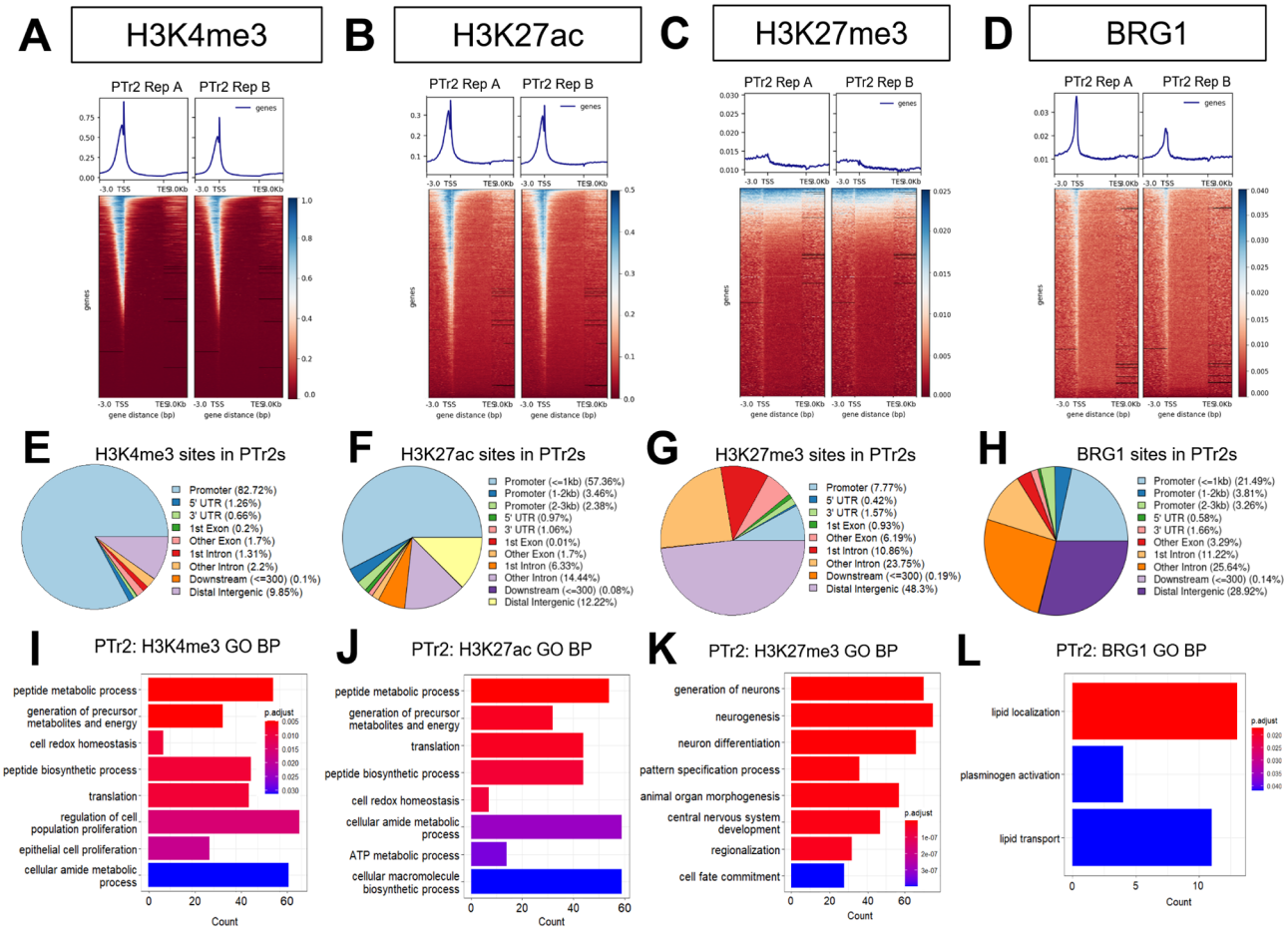

Fig. 2 Genomic localization and functional annotation of epigenetic features in PTr2 cells. Enrichment of (**A**) H3K4me3, (**B**) H3K27ac, (**C**) H3K27me3, and (**D**) BRG1 relative to global TSS and TES locations. Genomic feature annotation of global (**E**) H3K4me3, (**F**) H3K27ac, (**G**) H3K27me3, and (**H**) BRG1 binding. GO biological process overrepresentation analysis ($p < 0.05$ using one-sided Fisher's exact test; adjusted $P < 0.05$ after Benjamini-Hochberg FDR) for (**I**) H3K4me3, (**J**) H3K27ac, (**K**) H3K27me3, and (**L**) BRG1

of BRG1 signal was detected at global promoter regions, while 36.86% of signal was found in introns and 28.92% in distal intergenic regions (Fig. 2H). Once again, functional annotation of the global enrichment for these epigenetic features in PTr2 cells showed that a variety of different biological processes were enriched in the feature data, such as metabolic processes and control of cell proliferation (Fig. 2I-L). Similar to what was seen in PFF cells, the RNA-seq expression values for genes in the top 3 biological process categories for each histone mark showed that the vast majority of genes marked by H3K4me3 and H3K27ac were expressed, while the converse was seen for genes marked with H3K27me3 (Supplementary Fig. S3A-C).

Genome-wide chromatin state predictions in PFF and PTr2 cells

Using the alignment files for H3K4me3, H3K27ac, H3K27me3, and BRG1, we constructed a chromatin state prediction model for each cell type. A total of 10 different states were defined using a combination of the previously discussed enrichment details and combined signal annotation predictions for each feature to generate a genome-wide view of putative chromatin states in PFFs and PTr2s. The 10 distinct chromatin states defined here represented promoter regions (TSS Flanking, TSS Active, and Bivalent/poised TSS), putative enhancers (including predicted enhancer strength for Putative active and Putative weak), quiescent/lowly-transcribed regions (Quiescent(low)), and repressed genomic regions (Repressed, Weak repressed). Overall, while PFF and PTr2 cells were analyzed separately, the chromatin state predictions showed similar annotation patterns between the two cell lines (Fig. 2A-J; Supplementary Fig. S4A-F). As such, the order of predicted chromatin states was the same between these cells (Fig. 2E, J). The chromatin state predictions were in agreement with genome browser views of epigenetic feature enrichment and the accompanying RNA-seq data, including at select loci representing genes with known expression in both (*ITGB1*; confirmed by RT-qPCR*)* cell lines or only one cell line, as in the case of *KRT8* in PTr2 cells (confirmed by RT-qPCR) and *COL3A1* in the fibroblasts [50](Zoppi et al., 2004; Li et al. 2021) [51] (Fig. 3K).

To further explore how the chromatin state prediction approach could be leveraged to reveal insights into the gene expression regulatory landscape, we obtained the genomic coordinates for predicted enhancer regions in both cell lines and conducted a motif enrichment analysis to investigate transcription factor binding at these regions. PFF and PTr2 cells shared a common overrepresented motif in FOSL1::JUND (AP-1), and both cell lines had a TEAD motif, albeit corresponding to different members of the TEAD family. The remaining motifs showed no overlap in identity between the cell types.

Broad regions of H3K4me3 signal at predicted TSS regions mark essential development and tissue-specific genes

While the majority (around 87–88% in both cell lines) of regions in predicted TSS states (emission states 1 and 2) were 2 kb in length or less, a small subset of regions spanning 4 kb or more were identified in both PFF and PTr2 cells. In PFF cells, 464 regions were found to be spanning at least 4 kb in length, representing 1.70% of the putative TSS regions that were identified in this cell line (Fig. 4A), while in PTr2 cells, 593 regions spanning at least 4 kb were found, representing 1.80% of all predicted PTr2 TSS regions (Fig. 4B). These regions were strongly enriched in H3K4me3, and genome browser evaluation of their coordinates revealed a variety of genes marked with this broad H3K4me3 pattern (Fig. 4C). Biological functional annotation of these broad H3K4me3 regions in both cell lines revealed that they were strongly associated with genes involved in developmental processes (Fig. 4D, E; Supplementary Fig. S6A, B). Additionally, genes marked with broad H3K4me3 were expressed (Supplementary Fig. S6C, D). When these gene associations from the top five most overrepresented biological processes were compared to genes identified as either PFF or PTr2-specific (using a *t*-statistic ranked list approach on RNA-seq data), 52% of the genes in the PFF biological process list were present in the top 10% of PFF-specific genes (27/52), while this proportion was just over 25% in PTr2 cells (Fig. 4F). According to RNA-seq data, these genes were consistently transcriptionally active in both cell lines (Fig. 4G, H).

Discussion

The functional organization of chromatin serves not only to efficiently package the genome within the nucleus but also to enable the modulation of gene expression across different developmental time points and in response to various stimuli. Early development in mammals involves the precise and dynamic regulation of gene expression by epigenetic mechanisms, an essential group of processes that govern differentiation, lineage specification, and the preservation of identity across multiple cell divisions. Epigenetic landscape studies during early development can help us better understand how gene expression is regulated across different developmental time points, and these details can also be used for comparison in studies focused on evaluating aberrant developmental circumstances. However, despite gradual progress by FAANG and similar initiatives to remedy the relative dearth of epigenetic information available for domestic species, many gaps remain in our knowledge surrounding the epigenetic landscape in these species, especially regarding

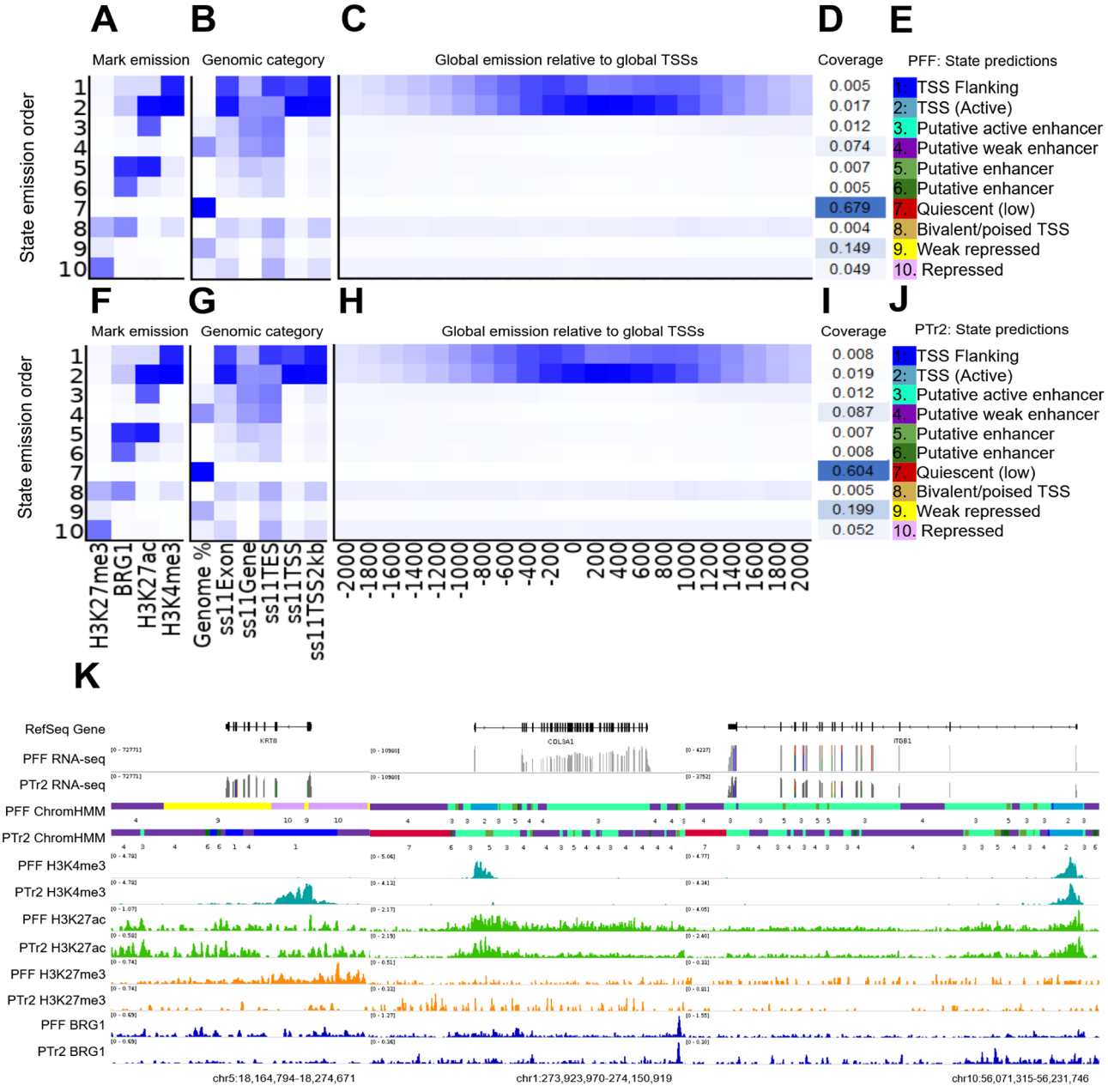

Fig. 3 Chromatin state prediction for PFF (**A-E**) and PTr2 (**F-J**) cells, (**A,F**) Emission probabilities for each profiled feature across 10 states. Darker blue indicates a higher emission probability (0–1), (**B, G**) Genomic coverage of each state relative to the defined genomic features. (**C, H**) Coverage of each state relative to global TSS coordinates. (**D,I**) Genomic coverage of each emission state relative to other states. (**E,J**) Predictions of chromatin state identities. (**K**) Genome browser view of PFF and PTr2 predicted chromatin states and associated epigenetic feature enrichment at select loci. Note that the color and numeric code in the predicted state tracks corresponds to the colors and numbers assigned in panels **E** and **J**

histone marks and chromatin state during early development. Continued efforts in this area are especially important given the immense combined value of the pig as a biomedical model and as a staple food animal species. The results of this study provide detailed enrichment characterizations of H3K4me3, H3K27ac, H3K27me3, and BRG1 in PFF and PTr2 cells, two porcine lines with many established applications in the fields of reproductive and molecular biology. Additionally, we further demonstrate that chromatin state prediction maps can be used to identify important epigenetic regulatory regions potentially involved in governing cell identity.

Our analysis of H3K4me3 enrichment in PFF and PTr2 cells revealed that the vast majority of H3K4me3 sites in both cell lines were in promoter regions around transcriptional start sites of transcriptionally active genes, an observation that is consistent with the character of this mark as reported in the literature. H3K4me3 is one of

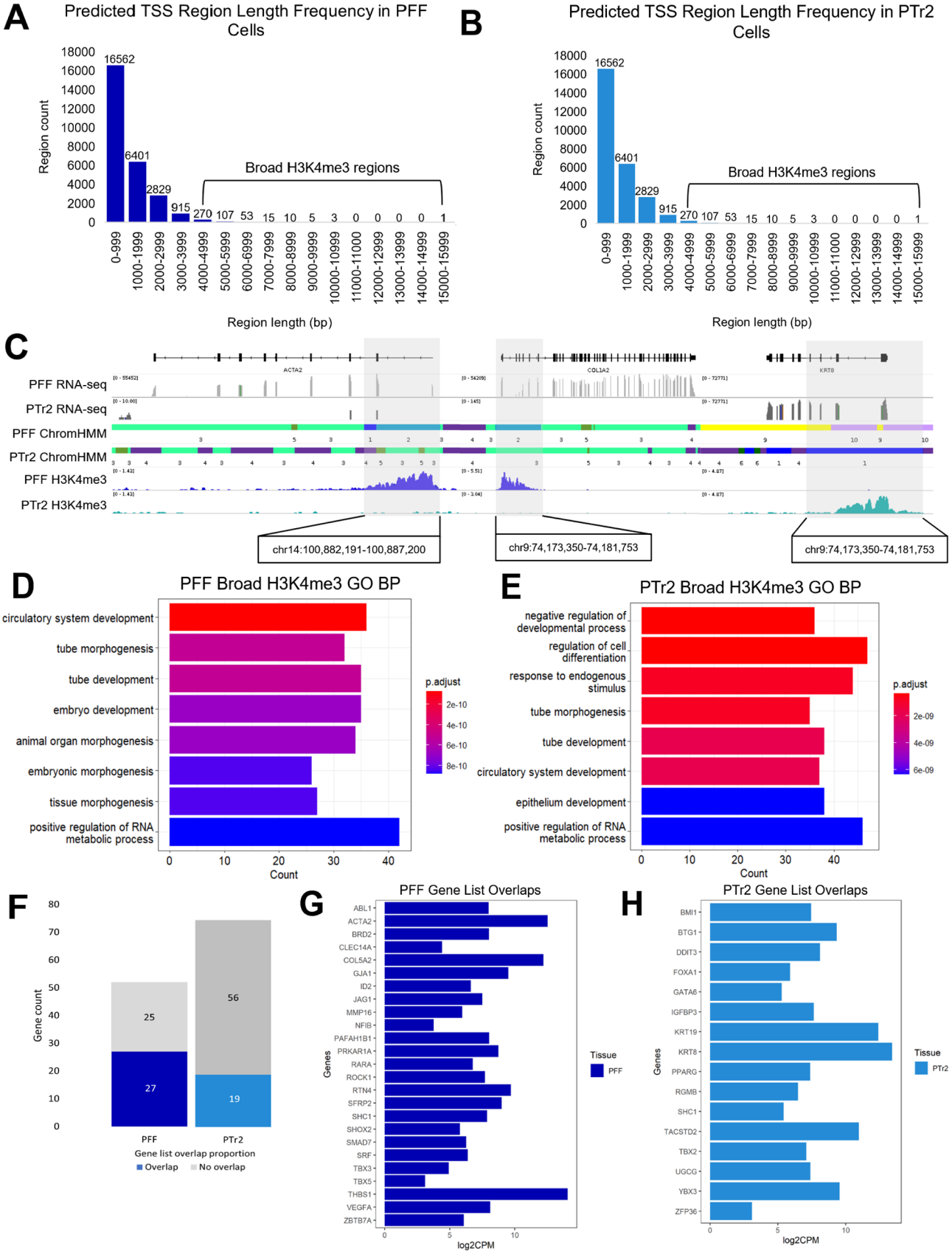

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Characterization of broad H3K4me3 mark regions and associated genes in PFF and PTr2 cells. Histogram of lengths of global predicted TSS regions in (**A**) PFF and (**B**) PTr2 cells. (**C**) Genome browser view of broad H3K4me3 enrichment at three representative loci in both cell lines. Biological process (BP) annotation for broad H3K4me3 regions in (**D**) PFF and (**E**) PTr2 cells ($p < 0.05$ using one-sided Fisher's exact test; adjusted $P < 0.05$ after Benjamini-Hochberg FDR). (**F**) Proportions of genes found both in the list of genes associated with the top 5 overrepresented BPs for broad H3K4me3 regions and in the top 10% of genes identified as cell-type specific using transcriptional data. Identities of genes overlapping the two list categories in (**G**) PFF and (**H**) PTr2 cells

the most widely profiled histone marks and is known to be enriched around active promoters during embryonic development and beyond [52–57]. Similarly, H3K27ac showed enrichment at promoter regions, though relatively less than H3K4me3, likely due to its presence at both promoters and enhancers [58, 59]. In both PFFs and PTr2s, several thousand (8,705 in PFFs and 7,416 in PTr2s) H3K4me3 and H3K27ac enrichment sites had overlapping genomic coordinates. The interplay between these PTMs has been investigated [60–62] and it has been posited that the upstream presence of H3K27ac may guide the installation of H3K4me3 at promoters by the acetylated histone reader BRD2 [62], though investigations into how H3K4me3 and H3K27ac may cooperate to influence transcription are ongoing.

H3K27me3, which is generally associated with transcriptional repression, showed lower enrichment at gene start and end sites compared to H3K4me3 and H3K27ac, mainly appearing in intergenic regions. Some active genes exhibited both H3K4me3 and H3K27me3 marks, possibly indicating bivalent regulation. The protein encoded by *FGF9*, the sole gene bound by both H3K4me3 and H3K27me3 in PFFs in the overrepresentation analysis, is known to have many roles in early development, including sex determination [63], cell proliferation [64], and morphogenesis [65–68]. Bivalent regulation of developmentally associated genes has been reported [69, 70], including *FGF9* [71–73], and while it is perhaps more commonly associated with pluripotency maintenance, it has been posited that this regulatory pattern may also play a role in modulating tissue-specific gene expression during later stages of development and into adulthood [69]. Indeed, in both cell lines, H3K27me3 was most enriched for developmentally associated terms relative to the other marks. Genome browser exploration of mark enrichment at these developmentally-related genes showed that some genes were bivalently marked with H3K4me3 and H3K27me3, while others were marked with only H3K27me3. RNA-seq data for these genes showed variable expression levels, and genes marked only with H3K27me3 were broadly not expressed according to this data. For example, *BMP4* was present in the list of developmentally-associated genes marked with H3K27me3 in PFF cells, but the corresponding RNA-seq data indicated that this gene was expressed, albeit at relatively low levels, in this cell type. *BMP4* is known to be implicated in a variety of developmental processes such as tissue morphogenesis [74–77]. Genome browser views of the *BMP4* locus indicate the enrichment of both H3K4me3 and H3K27me3, with H3K27me3 signal at and/or immediately upstream of gene promoter regions. In contrast, *KLK4*, an ameloblast-associated gene [78], was also in this list of H3K27me3-enriched genes, but it was not expressed based on RNA-seq data. Examination of the epigenetic status of this gene indicated that it was marked only by H3K27me3, consistent with its repressed character in this cell type.

BRG1 (for Brahma-related gene 1; encoded by SMARCA4) is one of two mutually exclusive central ATPases present within mammalian SWI/SNF (BAF) chromatin remodeling complexes, along with Brahma (BRM, encoded by SMARCA2). BRG1 contains a bromodomain that is capable of recognizing and binding acetylated histone lysine residues [79, 80], an event that results in SWI/SNF recruitment to genomic regions. Our analysis of global BRG1 enrichment in PFFs and PTr2s revealed that while some BRG1 presence could be detected at promoters, the vast majority of this remodeling enzyme's signal was detected outside of promoter regions, specifically in introns and distal intergenic stretches. RNA-seq data indicated that BRG1 was present at or around some transcriptionally active genes marked with H3K4me3, but not always consistently. BRG1 localization to promoter regions has been shown to initiate transcription following SWI/SNF-dependent chromatin remodeling [81–83]. There is also evidence to suggest that BRG1 can function as a transcriptional silencer [84] or regulator of H3K4me3 levels at promoters [85], highlighting the capacity of the SWI/SNF chromatin remodeling complex to exert versatile transcriptional control. BRG1 can bind both promoters and enhancers, with aberrant BAF complex activity affecting enhancer function and bivalent promoter regulation in cancers [86–93]. Our results indicated that a large number of transcriptionally active genes involved in the regulation of cell population proliferation and organism development were enriched with both BRG1 and H3K27ac in PFFs, which may suggest that chromatin accessibility at these genes is partly regulated through a mechanism facilitated by the recognition of H3K27ac by the BRG1 subunit of SWI/SNF complexes, though further examination of this relationship with additional chromatin profiling techniques would be needed to provide any mechanistic conclusions. The lack of genes identified to be co-bound by BRG1 and

H3K27ac in PTr2 cells could be due, in part, to the stringency of the applied statistical parameters, but it may also be at least partly reflective of a relatively lower frequency of BRG1/H3K27ac co-occupancy in this particular cell line.

The chromatin state prediction maps generated using these epigenetic features showed widespread conservation of certain regulatory regions such as TSSs between the two cell lines, except at cell type-specific genes. Additionally, transcription factor motif analysis of putative enhancer elements showed some similarities in motif conservation between the two cell lines, including the FOSL/JUND heterodimer, and TEAD family transcription factors, each of which have several known associations with regulating fundamental early developmental processes [94–97]. The transcription factor motifs that differed between PFF and PTr2 cells still included transcription factors with established implications in developmental processes such as cell proliferation, differentiation, and cell fate determination, suggesting that these differences may be involved in the coordination of cell type-specific gene expression programs. The chromatin state maps also depicted discernable differences in TSS/promoter region lengths and facilitated the identification of a subset of genomic regions marked with broad (>4 kb) H3K4me3 signal. Most H3K4me3 enrichment is 1–2 kb in length immediately upstream and towards the 5' end of gene bodies, though a subset of broad H3K4me3 enrichment patterns ranging from around 4 kb to upwards of 60 kb in length has been reported and associated with maintenance of cell identity [98–100]. Our findings in PFF and PTR2 cells support the observations that broad H4K3me3 domains may be implicated in regulating the expression of cell type-specific genes, as these subsets were strongly associated with genes involved in essential developmental processes, including transcriptionally active genes identified as being in the top 10% of cell type-specific genes in each cell line. Analysis of RNA-seq data indicated that genes marked by broad H3K4me3 in PFF and PTr2 cells are expressed, suggesting that broad H3K4me3 may have an activating role in these cell types. Broad H3K4me3 is reportedly associated with consistent and relatively high gene expression levels in many cell types [98, 101–103]. However, there is also evidence to suggest that broad H3K4me3 may have more repressive roles in certain contexts, such as in early embryo development, where broad H3K4me3 regions must be removed in order for zygotic genome activation (ZGA) to proceed [57, 104–106]. Given the need for precise temporal and spatial regulation of gene expression and the dynamic nature of chromatin organization during early embryonic development, the broad H3K4me3 domain may serve as an example of how epigenetic mechanisms exert their regulatory functions in a context-dependent manner.

In a 2021 paper, Pan and colleagues created chromatin state prediction maps across 14 different porcine tissues as part of a broader effort to construct the beginnings of an atlas of functional genetic elements in this species [21]. Part of this research was to bring attention to efforts by the Functional Annotation of Animal Genomes (FAANG) Consortium to bring together datasets produced by RNA-seq, chromatin profiling, DNA methylation studies, and chromatin interactions and accessibility in domestic livestock to address the relative lack of genomic annotation in these species [107]. While the chromatin state details from the present work are not as comprehensive and informative as the aforementioned study (which combined ChIP-seq, ATAC-seq, Reduced-representation bisulfite sequencing (RRBS-Seq), and RNA-seq), they do represent two cell types and developmental time points not presently captured in the atlas. Indeed, the authors of the Pan et al. study conclude that much greater quantities of epigenomic data will need to be collected across more developmental stages and tissues, particularly reproductive tissues, to develop a deeper understanding of the genotype-to-phenotype relationship and the complex underlying mechanisms and functionalities that influence development and disease.

Conclusions

Taken together, these findings provide a view of the epigenetic landscape and predicted chromatin state in PFF and PTr2 cells, highlighting both the similarities and differences in epigenetic feature and regulatory region localization involved in maintaining essential functions and conferring unique cell identity. As pigs continue to grow in popularity as biomedical models for human-oriented research, likely, concerted efforts to understand the mechanistic underpinnings guiding gene expression in this species will continue to be examined and compared to human contexts in order to maximize the utility and efficiency of using swine as a model organism for translational research.

Abbreviations

BRG1	Brahma-related gene 1
CUT&RUN	Cleavage under targets & release using nuclease
EZH2	Enhancer of zeste homolog 2
FGFR2	Fibroblast growth factor receptor 2
FAANG	Functional Annotation of Animal Genomes
GO	Gene Ontology
KGFR	Keratinocyte growth factor receptor
NGS	Next-generation sequencing
pAG-MNase	Protein A/G micrococcal nuclease
PFF	Porcine fetal fibroblast
PTM	Post-translational modification
PTr2	Porcine trophectoderm
RRBS-seq	Reduced-representation bisulfite sequencing
SCNT	Somatic cell nuclear transfer
TES	Transcriptional end site

TSS	Transcriptional start site
ZGA	Zygotic genome activation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13072-024-00542-w.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Acknowledgements
The authors wish to thank Dr. Laurie Jaeger for providing the PTr2 cells used in this work. Additionally, we wish to acknowledge and thank the Purdue Genomics Core for services related to sample and library quality control and Novogene for providing sequencing services.

Author contributions
S.I. prepared the CUT&RUN samples, analyzed and interpreted the data, prepared the figures, and drafted the manuscript. R.C. reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding
This project was supported, in part, by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (award number R01HD084309). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability
The CUT&RUN datasets generated and analyzed during the current study are available in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1036691. The RNA-seq data analyzed during the current study are available for download from the NCBI SRA BioProjects PRJNA798047 and PRJNA778857.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Animal Sciences, Purdue University, West Lafayette, IN 47907, USA

Received: 9 November 2023 / Accepted: 16 May 2024
Published online: 21 May 2024

References

1. Sykes M. Developing pig-to-human organ transplants. Science. 2022;378(6616):135–6.
2. Xi J, Zheng W, Chen M, Zou Q, Tang C, Zhou X. Genetically engineered pigs for xenotransplantation: hopes and challenges. Front Cell Dev Biol. 2023;10:1093534.
3. Lai L, Park KW, Cheong HT, Kühholzer B, Samuel M, Bonk A, et al. Transgenic pig expressing the enhanced green fluorescent protein produced by nuclear transfer using colchicine-treated fibroblasts as donor cells. Mol Reprod Dev. 2002;62(3):300–6.
4. Hyun S, Lee G, Kim D, Kim H, Lee S, Nam D, et al. Production of Nuclear transfer-derived piglets using porcine fetal fibroblasts transfected with the enhanced green fluorescent Protein1. Biol Reprod. 2003;69(3):1060–8.
5. Aigner B, Renner S, Kessler B, Klymiuk N, Kurome M, Wünsch A, et al. Transgenic pigs as models for translational biomedical research. J Mol Med. 2010;88(7):653–64.
6. Renner S, Fehlings C, Herbach N, Hofmann A, von Waldthausen DC, Kessler B, et al. Glucose intolerance and reduced proliferation of pancreatic β-Cells in transgenic pigs with impaired glucose-dependent insulinotropic polypeptide function. Diabetes. 2010;59(5):1228–38.
7. Stoltz DA, Meyerholz DK, Pezzulo AA, Ramachandran S, Rogan MP, Davis GJ, et al. Cystic fibrosis pigs develop lung disease and exhibit defective bacterial eradication at Birth. Sci Transl Med. 2010;2(29):29ra31.
8. Meurens F, Summerfield A, Nauwynck H, Saif L, Gerdts V. The pig: a model for human infectious diseases. Trends Microbiol. 2012;20(1):50–7.
9. Fang B, Ren X, Wang Y, Li Z, Zhao L, Zhang M, et al. Apolipoprotein E deficiency accelerates atherosclerosis development in miniature pigs. Dis Model Mech. 2018;11(10):dmm036632.
10. Ju H, Zhang J, Bai L, Mu Y, Du Y, Yang W, et al. The transgenic cloned pig population with integrated and controllable GH expression that has higher feed efficiency and meat production. Sci Rep. 2015;5:10152.
11. Zhang X, Li Z, Yang H, Liu D, Cai G, Li G, et al. Novel transgenic pigs with enhanced growth and reduced environmental impact. eLife. 2018;7:e34286.
12. Yuan H, Yang L, Zhang Y, Xiao W, Wang Z, Tang X, et al. Current status of genetically modified pigs that are resistant to virus infection. Viruses. 2022;14(2):417.
13. Jaeger LA, Spiegel AK, Ing NH, Johnson GA, Bazer FW, Burghardt RC. Functional effects of transforming growth factor β on Adhesive properties of Porcine Trophectoderm. Endocrin. 2005;146(9):3933–42.
14. Kong X, Tan B, Yin Y, Gao H, Li X, Jaeger LA, et al. l-Arginine stimulates the mTOR signaling pathway and protein synthesis in porcine trophectoderm cells. J Nutr Biochem. 2012;23(9):1178–83.
15. Kong X, Wang X, Yin Y, Li X, Gao H, Bazer FW, et al. Putrescine stimulates the mTOR Signaling Pathway and protein synthesis in Porcine Trophectoderm Cells1. Biol Reprod. 2014;91(5):106.
16. Han D, Jiang L, Gu X, Huang S, Pang J, Wu Y, et al. SIRT3 deficiency is resistant to autophagy-dependent ferroptosis by inhibiting the AMPK/mTOR pathway and promoting GPX4 levels. J Cell Physio. 2020;235(11):8839–51.
17. Wang H, Liu Y, Zhou T, Gao L, Li J, Wu X et al. Uridine affects amino acid metabolism in the sow-piglets model and increases viability of pTr2 cells. Front Nutr. 2022;9.
18. Foissac S, Djebali S, Munyard K, Vialaneix N, Rau A, Muret K, et al. Multi-species annotation of transcriptome and chromatin structure in domesticated animals. BMC Biol. 2019;17(1):108.
19. Halstead MM, Kern C, Saelao P, Wang Y, Chanthavixay G, Medrano JF, et al. A comparative analysis of chromatin accessibility in cattle, pig, and mouse tissues. BMC Genomics. 2020;21(1):698.
20. Kern C, Wang Y, Xu X, Pan Z, Halstead M, Chanthavixay G, et al. Functional annotations of three domestic animal genomes provide vital resources for comparative and agricultural research. Nat Commun. 2021;12(1):1821.
21. Pan Z, Yao Y, Yin H, Cai Z, Wang Y, Bai L, et al. Pig genome functional annotation enhances the biological interpretation of complex traits and human disease. Nat Commun. 2021;12(1):5848.
22. Andersson L, Archibald AL, Bottema CD, Brauning R, Burgess SC, Burt DW, et al. Coordinated international action to accelerate genome-to-phenome with FAANG, the functional annotation of animal genomes project. Genome Biol. 2015;16(1):57.
23. Kühholzer B, Hawley RJ, Lai L, Kolber-Simonds D, Prather RS. Clonal lines of transgenic fibroblast cells derived from the same fetus result in different development when used for nuclear transfer in pigs. Biol Reprod. 2001;64(6):1695–8.
24. Ka H, Jaeger LA, Johnson GA, Spencer TE, Bazer FW. Keratinocyte growth factor is Up-Regulated by Estrogen in the Porcine Uterine Endometrium and functions in Trophectoderm Cell Proliferation and differentiation. Endocrin. 2001;142(6):2303–10.

25. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat Protoc. 2018;13(5):1006–19.
26. Krueger F, James F, Ewels P, Afyounian E, Schuster-Boeckler B. FelixKrueger/TrimGalore: v0.6.7 - DOI via Zenodo (0.6.7). 2021; https://doi.org/10.5281/zenodo.5127899.
27. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
28. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2):giab008.
29. Tay RE, Olawoyin O, Cejas P, Xie Y, Meyer CA, Ito Y, et al. Hdac3 is an epigenetic inhibitor of the cytotoxicity program in CD8 T cells. J Exp Med. 2020;217(7):e20191453.
30. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next-generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44:W160–165.
31. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integr Genomics Viewer Nat Biotechnol. 2011;29(1):24–6.
32. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, et al. The UCSC table browser data retrieval tool. Nucleic Acids Res. 2004;32:D493–496.
33. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
34. Stark R, Brown G. DiffBind: differential binding analysis of ChIP-Seq peak data. 2011;http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf.
35. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcomes in breast cancer. Nature. 2012;481(7381):389–93.
36. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31(14):2382–3.
37. Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinform. 2010;11(1):237.
38. Ernst J, Kellis M. Chromatin state discovery and genome annotation with ChromHMM. Nat Protoc. 2017;12(12):2478–92.
39. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
40. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinform. 2015;31(2):166–9.
41. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
42. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):R25.
43. Finucane HK, Reshef YA, Anttila V, Slowikowski K, Gusev A, Byrnes A, et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nat Genet. 2018;50(4):621–9.
44. Fang L, Cai W, Liu S, Canela-Xandri O, Gao Y, Jiang J, et al. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. Genome Res. 2020;30(5):790–801.
45. Jaeger LA, Spiegel AK, Ing NH, Johnson GA, Bazer FW, Burghardt RC. Functional effects of transforming growth factor β on Adhesive properties of Porcine Trophectoderm. Endocrinology. 2005;146(9):3933–42.
46. Liu B, Paudel S, Flowers WL, Piedrahita JA, Wang X. Uterine histotroph and conceptus development: III. Adrenomedullin stimulates proliferation, migration and adhesion of porcine trophectoderm cells via AKT-TSC2-MTOR cell signaling pathway. Amino Acids. 2023;55(6):743–56.
47. Ka H, Jaeger LA, Johnson GA, Spencer TE, Bazer FW. Keratinocyte growth factor is Up-Regulated by Estrogen in the Porcine Uterine Endometrium and functions in Trophectoderm Cell Proliferation and Differentiation*. Endocrinology. 2001;142(6):2303–10.
48. Erikson DW, Burghardt RC, Bayless KJ, Johnson GA. Secreted phosphoprotein 1 (SPP1, osteopontin) binds to integrin Alphavbeta6 on Porcine Trophectoderm cells and integrin Alphavbeta3 on uterine luminal epithelial cells, and promotes Trophectoderm Cell Adhesion and Migration1. Biol Reprod. 2009;81(5):814–25.
49. Frank JW, Seo H, Burghardt RC, Bayless KJ, Johnson GA. ITGAV (alpha v integrins) bind SPP1 (osteopontin) to support trophoblast cell adhesion. Reproduction. 2017;153(5):695–706.
50. Zoppi N, Gardella R, Paepe AD, Barlati S, Colombi M. Human fibroblasts with mutations in COL5A1 and COL3A1 genes do not organize Collagens and Fibronectin in the Extracellular Matrix, down-regulate α2β1 integrin, and Recruit αvβ3 instead of α5β1 integrin *. J Biol Chem. 2004;279(18):18157–68.
51. Li W, Chi N, Rathnayake RAC, Wang R. Distinctive roles of fibrillar collagen I and collagen III in mediating fibroblast-matrix interaction: a nanoscopic study. Biochem Biophys Res Commun. 2021;560:66–71.
52. Hon GC, Hawkins RD, Ren B. Predictive chromatin signatures in the mammalian genome. Hum Mol Genet. 2009;18(R2):R195–201.
53. Grandy RA, Whitfield TW, Wu H, Fitzgerald MP, VanOudenhove JJ, Zaidi SK, et al. Genome-wide studies reveal that H3K4me3 modification in bivalent genes is dynamically regulated during the pluripotent cell cycle and stabilized upon differentiation. Mol Cell Biol. 2016;36(4):615–27.
54. Liu X, Wang C, Liu W, Li J, Li C, Kou X, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature. 2016;537(7621):558–62.
55. Hsu CL, Lo YC, Kao CF. H3K4 methylation in aging and metabolism. Epigenomes. 2021;5(2):14.
56. Lambrot R, Chan D, Shao X, Aarabi M, Kwan T, Bourque G, et al. Whole-genome sequencing of H3K4me3 and DNA methylation in human sperm reveals regions of overlap linked to fertility and development. Cell Rep. 2021;36(3):109418.
57. Bu G, Zhu W, Liu X, Zhang J, Yu L, Zhou K, et al. Coordination of zygotic genome activation entry and exit by H3K4me3 and H3K27me3 in porcine early embryos. Genome Res. 2022;32(8):1487–501.
58. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009;459(7243):108–12.
59. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences. 2010;107(50):21931–6.
60. Tie F, Banerjee R, Saiakhova AR, Howard B, Monteith KE, Scacheri PC, et al. Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize polycomb silencing. Development. 2014;141(5):1129–39.
61. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications – writers that read. EMBO Rep. 2015;16(11):1467–81.
62. Zhao W, Xu Y, Wang Y, Gao D, King J, Xu Y, et al. Investigating crosstalk between H3K27 acetylation and H3K4 trimethylation in CRISPR/dCas-based epigenome editing and gene activation. Sci Rep. 2021;11(1):15912.
63. Sánchez L, Chaouiya C. Primary sex determination of placental mammals: a modelling study uncovers dynamical developmental constraints in the formation of sertoli and granulosa cells. BMC Syst Biol. 2016;10:37.
64. Chang MM, Wu SZ, Yang SH, Wu CC, Wang CY, Huang BM. FGF9/FGFR1 promotes cell proliferation, epithelial-mesenchymal transition, M2 macrophage infiltration and liver metastasis of lung cancer. Transl Oncol. 2021;14(11):101208.
65. Santos-Ocampo S, Colvin JS, Chellaiah A, Ornitz DM. Expression and biological activity of mouse fibroblast growth Factor-9 (*). J Biol Chem. 1996;271(3):1726–31.
66. Pirvola U, Zhang X, Mantela J, Ornitz DM, Ylikoski J. Fgf9 signaling regulates inner ear morphogenesis through epithelial-mesenchymal interactions. Dev Biol. 2004;273(2):350–60.
67. Yin Y, Ornitz DM. FGF9 and FGF10 activate distinct signaling pathways to direct lung epithelial specification and branching. Sci Signal. 2020;13(621):eaay4353.
68. Zhang X, Weng M, Chen Z. Fibroblast growth factor 9 (FGF9) negatively regulates the early stage of chondrogenic differentiation. PLoS ONE. 2021;16(2):e0241281.
69. Blanco E, González-Ramírez M, Alcaine-Colet A, Aranda S, Di Croce L. The Bivalent Genome: characterization, structure, and Regulation. Trends Genet. 2020;36(2):118–31.
70. Yu Y, Li X, Jiao R, Lu Y, Jiang X, Li X. H3K27me3-H3K4me1 transition at bivalent promoters instructs lineage specification in development. Cell Biosci. 2023;13:66.
71. Kurimoto K, Yabuta Y, Hayashi K, Ohta H, Kiyonari H, Mitani T, et al. Quantitative Dynamics of Chromatin Remodeling during germ cell specification from mouse embryonic stem cells. Cell Stem Cell. 2015;16(5):517–32.
72. Taube JH, Sphyris N, Johnson KS, Reisenauer KN, Nesbit TA, Joseph R, et al. The H3K27me3-demethylase KDM6A is suppressed in breast cancer stem-like cells, and enables the resolution of bivalency during the mesenchymal-epithelial transition. Oncotarget. 2017;8(39):65548–65.

73. Brennan K, Zheng H, Fahrner JA, Shin JH, Gentles AJ, Schaefer B, et al. NSD1 mutations deregulate transcription and DNA methylation of bivalent developmental genes in Sotos syndrome. Hum Mol Genet. 2022;31(13):2164–84.
74. Watanabe Y, Le Douarin NM. A role for BMP-4 in the development of subcutaneous cartilage. Mech Dev 57:69–78.
75. Hogan BL. Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes Dev. 1996;10(13):1580–94.
76. Botchkarev VA, Sharov AA. BMP signaling in the control of skin development and hair follicle growth. Differentiation. 2004;72(9–10):512–26.
77. Zhang P, Li J, Tan Z, Wang C, Liu T, Chen L, et al. Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. Blood. 2008;111(4):1933–41.
78. Lu Y, Papagerakis P, Yamakoshi Y, Hu JCC, Bartlett JD, Simmer JP. Functions of KLK4 and MMP-20 in dental enamel formation. Biol Chem. 2008;389(6):695–700.
79. Haynes SR, Dollard C, Winston F, Beck S, Trowsdale J, Dawid IB. The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. Nucleic Acids Res. 1992;20(10):2603.
80. Tang L, Nogales E, Ciferri C. Structure and function of SWI/SNF Chromatin Remodeling Complexes and mechanistic implications for transcription. Prog Biophys Mol Biol. 2010;102(2–3):122–8.
81. King HW, Klose RJ. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. Davidson I, editor. eLife. 2017;6:e22631.
82. Chen Z, Gao Y, Yao L, Liu Y, Huang L, Yan Z, et al. LncFZD6 initiates Wnt/β-catenin and liver TIC self-renewal through BRG1-mediated FZD6 transcriptional activation. Oncogene. 2018;37(23):3098–112.
83. Zhou Q, Zhang Y, Wang B, Zhou W, Bi Y, Huai W, et al. KDM2B promotes IL-6 production and inflammatory responses through Brg1-mediated chromatin remodeling. Cell Mol Immunol. 2020;17(8):834–42.
84. Ma P, Pan Y, Yang F, Fang Y, Liu W, Zhao C, et al. KLF5-Modulated lncRNA NEAT1 contributes to Tumorigenesis by acting as a Scaffold for BRG1 to silence GADD45A in gastric Cancer. Mol Therapy - Nucleic Acids. 2020;22:382–95.
85. Xiao C, Hou J, Wang F, Song Y, Zheng J, Luo L, et al. Endothelial Brg1 fine-tunes notch signaling during zebrafish heart regeneration. Npj Regen Med. 2023;8(1):1–13.
86. DiRenzo J, Shang Y, Phelan M, Sif S, Myers M, Kingston R, et al. BRG-1 is recruited to estrogen-responsive promoters and cooperates with factors involved in histone acetylation. Mol Cell Biol. 2000;20(20):7541–9.
87. Tolstorukov MY, Sansam CG, Lu P, Koellhoffer EC, Helming KC, Alver BH et al. Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. Proceedings of the National Academy of Sciences. 2013;110(25):10165–70.
88. Alexander JM, Hota SK, He D, Thomas S, Ho L, Pennacchio LA, et al. Brg1 modulates enhancer activation in mesoderm lineage commitment. Development. 2015;142(8):1418–30.
89. Mathur R, Alver BH, San Roman AK, Wilson BG, Wang X, Agoston AT, et al. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. Nat Genet. 2017;49(2):296–302.
90. Hodges HC, Stanton BZ, Cermakova K, Chang CY, Miller EL, Kirkland JG, et al. Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. Nat Struct Mol Biol. 2018;25(1):61–72.
91. Laubscher D, Gryder BE, Sunkel BD, Andresson T, Wachtel M, Das S, et al. BAF complexes drive proliferation and block myogenic differentiation in fusion-positive rhabdomyosarcoma. Nat Commun. 2021;12(1):6924.
92. Stanton BZ, Hodges C, Calarco JP, Braun SMG, Ku WL, Kadoch C, et al. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. Nat Genet. 2017;49(2):282–8.
93. Kadoch C, Williams RT, Calarco JP, Miller EL, Weber CM, Braun SMG, et al. Dynamics of BAF–Polycomb complex opposition on heterochromatin in normal and oncogenic states. Nat Genet. 2017;49(2):213–22.
94. Renaud SJ, Kubota K, Rumi MAK, Soares MJ. The FOS transcription factor family differentially controls trophoblast Migration and Invasion*. J Biol Chem. 2014;289(8):5025–39.
95. Sobolev VV, Khashukoeva AZ, Evina OE, Geppe NA, Chebysheva SN, Korsunskaya IM, et al. Role of the transcription factor FOSL1 in Organ Development and Tumorigenesis. Int J Mol Sci. 2022;23(3):1521.
96. Lin KC, Park HW, Guan KL. Regulation of the Hippo pathway transcription factor TEAD. Trends Biochem Sci. 2017;42(11):862–72.
97. Currey L, Thor S, Piper M. TEAD family transcription factors in development and disease. Development. 2021;148(12):dev196675.
98. Chen K, Chen Z, Wu D, Zhang L, Lin X, Su J, et al. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet. 2015;47(10):1149–57.
99. Beacon TH, Delcuve GP, López C, Nardocci G, Kovalchuk I, van Wijnen AJ, et al. The dynamic broad epigenetic (H3K4me3, H3K27ac) domain as a mark of essential genes. Clin Epigenetics. 2021;13(1):138.
100. Kent D, Marchetti L, Mikulasova A, Russell LJ, Rico D. Broad H3K4me3 domains: maintaining cellular identity and their implication in super-enhancer hijacking. BioEssays. 2023;45(10):e2200239.
101. Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, et al. H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell. 2014;158(3):673–88.
102. Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, et al. H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell. 2015;163(5):1281–6.
103. Belhocine M, Simonin M, Abad Flores JD, Cieslak A, Manosalva I, Pradel L, et al. Dynamics of broad H3K4me3 domains uncover an epigenetic switch between cell identity and cancer-related genes. Genome Res. 2022;32(7):1328–42.
104. Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature. 2016;537(7621):548–52.
105. Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature. 2016;537(7621):553–7.
106. Zhou C, Halstead MM, Bonnet-Garnier A, Schultz RM, Ross PJ. Histone remodeling reflects conserved mechanisms of bovine and human preimplantation development. EMBO Rep. 2023;24(3):e55726.
107. Giuffra E, Tuggle CK. Functional annotation of animal genomes (FAANG): current achievements and Roadmap. Annu Rev Anim Biosci. 2019;7(1):65–88.

Publisher's Note