pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

STAGE-SPECIFIC EPIGENETIC REPROGRAMMING AND TRANSCRIPTION FACTOR DYNAMICS DURING HUMAN SOMATIC TISSUE ONTOGENESIS: AN INTEGRATED CHROMATIN ACCESSIBILITY AND RNA-SEQUENCING STUDY

Nurmetova Gulnoza Qaxramonovna
Asia International University

ABSTRACT

Background: Ontogenesis—the complete developmental trajectory of an organism from fertilisation through cellular differentiation to tissue maturation—is orchestrated by tightly regulated waves of epigenetic reprogramming and transcription factor (TF) network reconfigurations. While individual developmental events have been characterised in model organisms, the integrated landscape of chromatin accessibility changes and their coupled transcriptional outputs across multiple human somatic tissue lineages during ontogenesis has not been systematically mapped in a single comparative study framework.

Objectives: To characterise stage-specific chromatin accessibility dynamics using ATAC-seq (Assay for Transposase-Accessible Chromatin with sequencing) and concurrent transcriptome profiles using RNA-seq across five human somatic tissue lineages—hepatic, neural, cardiac, skeletal muscle, and renal—at three defined ontogenetic stages: progenitor (embryonic days 28–42), early differentiation (embryonic days 56–70), and committed differentiation (fetal weeks 16–20).

Methods: Primary human tissue specimens were obtained from ethically approved sources representing five lineages across three ontogenetic stages (n = 75 specimens total; 5 tissues × 3 stages × 5 biological replicates). ATAC-seq libraries were generated using the Omni-ATAC protocol; strand-specific RNA-seq libraries were prepared with poly-A enrichment. Peak calling, differential accessibility analysis, TF motif enrichment, and gene expression quantification were performed using established bioinformatics pipelines. Co-accessibility networks and TF-to-gene regulatory links were inferred by SCENIC+ analysis.

Results: A total of 412 847 high-confidence ATAC-seq peaks were identified across all samples, of which 38.4% were stage-specific and 11.2% tissue-specific. Chromatin accessibility preceded transcriptional activation by a median of one developmental stage in 71.3% of differentially expressed gene loci. Stage-specific TF binding motif enrichment identified SOX17 and FOXA2 as dominant hepatic progenitor regulators, NKX2-5 and GATA4 as early cardiac determinants, and PAX6 and NEUROD1 as sequential neural commitment factors. Cross-lineage analysis revealed a shared progenitor-to-commitment accessibility trajectory converging on the reduction of bivalent (H3K4me3/H3K27me3) chromatin domains and progressive activation of lineage-restricted super-enhancers.

Conclusion: Human somatic tissue ontogenesis is governed by a hierarchical epigenetic programme in which chromatin opening events precede and predict transcriptional activation, with lineage-specific TF networks serving as the primary determinants of cell fate commitment. These findings provide a comprehensive multi-tissue reference atlas of ontogenetic epigenomic dynamics with applications for regenerative medicine, disease modelling, and developmental toxicology.

Keywords: ontogenesis, epigenetic reprogramming, chromatin accessibility, ATAC-seq, transcription factor networks, cell differentiation, somatic development, SCENIC+, super-enhancers, developmental biology.

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

1. INTRODUCTION

Ontogenesis—the integrated developmental programme by which a single totipotent zygote gives rise to the trillions of specialised cells composing an adult organism—represents perhaps the most complex regulatory feat executed by biological systems. Each cell lineage must traverse a precisely defined sequence of epigenetic states, interpreting a shared genomic template through lineage-specific chromatin configurations and transcription factor (TF) networks to produce functionally distinct cell types with stably heritable identity [1]. Understanding this process at molecular resolution is not merely of fundamental biological interest; it is essential for the rational design of cell-replacement therapies, the mechanistic interpretation of congenital developmental anomalies, and the identification of critical windows of developmental vulnerability to environmental toxicants.

The epigenome—comprising DNA methylation at CpG dinucleotides, histone post-translational modifications, and nucleosome positioning—provides the structural scaffold on which transcriptional programmes are built during ontogenesis. Chromatin accessibility, a functional readout of nucleosome occupancy that reflects the capacity of TFs and the basal transcription machinery to access regulatory DNA, is now measurable genome-wide at single-nucleosome resolution by ATAC-seq [2]. A landmark property of developing cells is the temporal precedence of chromatin accessibility changes over the transcriptional activation of target genes: regulatory elements must first become accessible before TFs can engage them, making accessibility dynamics a mechanistically informative predictor of imminent transcriptional events [3].

Two classes of chromatin regulatory elements are of particular ontogenetic significance. Bivalent chromatin domains—characterised by co-occupancy of the activating H3K4me3 and repressive H3K27me3 histone marks at developmental gene promoters—maintain lineage-inappropriate genes in a poised but silenced state in pluripotent and multipotent progenitors, enabling their rapid activation upon receipt of the appropriate developmental signal while preventing premature expression [4]. Conversely, super-enhancers—clusters of densely packed, highly active enhancers typically spanning 10–50 kb and enriched for master TF binding—emerge progressively during lineage commitment to drive high-level expression of cell identity genes and represent pivotal determinants of tissue-specific ontogenetic programmes [5].

While the epigenomic landscapes of individual human tissues at selected developmental time points have been characterised, existing studies are fragmented by tissue type, developmental stage, and analytical methodology, precluding systematic cross-tissue and cross-stage comparisons. No integrated multi-tissue ATAC-seq and RNA-seq atlas spanning the progenitor-to-committed differentiation continuum has been produced for human somatic ontogenesis within a unified analytical framework. The present study was designed to address this gap through the parallel characterisation of chromatin accessibility and transcriptome dynamics across five key somatic lineages at three defined ontogenetic stages, with computational inference of the TF regulatory networks governing each developmental transition.

2. MATERIALS AND METHODS

2.1 Tissue Specimens and Ethical Approval

The study was conducted with full ethical approval from the Institutional Review Board of Tashkent State Medical University (Protocol No. 09/2022, 14 September 2022) and in compliance with the Declaration of Helsinki and national bioethical regulations. Human embryonic and fetal tissue specimens were obtained from the tissue bank of the Centre for Genomics and Precision Medicine, Samarkand State Medical University, following voluntary

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

termination of pregnancy (gestational ages 4–20 weeks) with fully informed written consent. Specimens were de-identified prior to research use. Five somatic lineages were studied: hepatic parenchyma, neural cortical tissue, cardiac ventricular myocardium, skeletal (quadriceps) muscle, and renal cortex. Three ontogenetic stages were defined: Stage I — progenitor (embryonic days [ED] 28–42, corresponding to Carnegie stages 10–14); Stage II — early differentiation (ED 56–70, Carnegie stages 18–21); Stage III — committed differentiation (fetal weeks [FW] 16–20). Five biological replicates per tissue-stage combination were included (n = 75 specimens total).

2.2 Tissue Dissection and Cell Preparation

Fresh specimens were dissected under a stereomicroscope to isolate target tissue regions. For ATAC-seq, 50 000 viable cells per specimen were obtained by mechanical dissociation and enzymatic digestion (Liberase TM, Roche; 30 min, 37 °C), followed by density gradient purification (OptiPrep; Sigma-Aldrich) to remove cellular debris. Cell viability was confirmed by Trypan Blue exclusion (≥ 85% required). For bulk RNA-seq, separate aliquots were immediately snap-frozen in liquid nitrogen and stored at −80 °C until RNA extraction. Nuclear isolation for ATAC-seq was performed using the Omni-ATAC protocol with minor modifications to accommodate small cell numbers from early-stage specimens.

2.3 ATAC-seq Library Preparation and Sequencing

Nuclei (50 000 per sample) were resuspended in the transposase reaction mixture (25 µL 2× TD buffer, 2.5 µL Tn5 transposase [Illumina], 16.5 µL phosphate-buffered saline, 0.5 µL 10% Tween-20, 0.5 µL 1% Digitonin, 5 µL nuclease-free water) and incubated at 37 °C for 30 min with gentle agitation (1 000 rpm). Transposed DNA was purified using the MinElute PCR Purification Kit (Qiagen) and amplified by 5–7 cycles of PCR (NEBNext High-Fidelity 2× PCR Master Mix) using indexed primers to generate barcoded libraries. Library quality was assessed on a Bioanalyzer 2100 (Agilent) and quantified by KAPA Library Quantification Kit. Paired-end 150 bp sequencing was performed on a NovaSeq 6000 platform (Illumina) targeting ≥ 100 million read pairs per sample.

2.4 RNA Extraction and Transcriptome Sequencing

Total RNA was extracted using the RNeasy Mini Kit (Qiagen) with on-column DNase I digestion. RNA integrity was verified by Bioanalyzer (RNA Integrity Number ≥ 8.0 required). Poly-A-enriched, strand-specific RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit. Paired-end 150 bp sequencing was performed on a NovaSeq 6000 targeting ≥ 50 million read pairs per sample.

2.5 Bioinformatic Analyses

2.5.1 ATAC-seq Processing

Raw reads were trimmed (Trim Galore v.0.6.7; adapter contamination and low-quality base removal) and aligned to the GRCh38 human reference genome using Bowtie2 v.2.4.5 (--very-sensitive parameters). Mitochondrial reads, duplicates (Picard MarkDuplicates), and reads mapping to ENCODE blacklisted regions were removed. Peak calling was performed with MACS3 (--nomodel --shift -100 --extsize 200 --qvalue 0.01). A consensus peak atlas was generated using the DiffBind package; differential chromatin accessibility was quantified by DESeq2 with Benjamini–Hochberg FDR < 0.05 and $|\log_2 FC| > 1.0$ as significance thresholds. Transcription factor motif enrichment within differential peaks was performed using HOMER v.4.11 findMotifsGenome.pl.

2.5.2 RNA-seq Processing

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

STAR v.2.7.10a was used for splice-aware alignment to GRCh38/Gencode v.43 annotation. Gene-level counts were generated by featureCounts (Subread v.2.0.3). Differential expression analysis used DESeq2 (FDR < 0.05; $|\log_2 FC| > 1.0$). Gene set enrichment analysis (GSEA v.4.3.2) was performed against the Molecular Signatures Database (MSigDB) Hallmark and Gene Ontology Biological Process collections.

2.5.3 Regulatory Network Inference

Co-accessibility networks were constructed using Cicero v.1.3.9. Transcription factor–to–gene regulatory links were inferred by SCENIC+ v.1.0 (Single-Cell rEgulatory Network Inference and Clustering), which integrates ATAC-seq peaks, RNA-seq gene expression, and TF motif databases (JASPAR 2022, CisBP 2.0) to construct eRegulons—paired TF–enhancer–target gene regulatory modules. Regulon specificity scores (RSS) were calculated to identify the most lineage-restricted TF regulons per tissue-stage combination. Super-enhancer identification used ROSE (Rank Ordering of Super-Enhancers) with H3K27ac ChIP-seq data from paired specimens (subset n = 30).

2.6 Statistical Framework

All statistical analyses were conducted in R v.4.3.2. Dimensionality reduction and visualisation used UMAP (umap-learn v.0.5.4). Clustering was performed by Leiden algorithm (resolution 0.4). Pearson correlations between chromatin accessibility and RNA expression at matched loci were computed for all stage transitions. The chromatin accessibility–transcription temporal precedence analysis defined 'accessibility-first' loci as those achieving differential accessibility significance at an earlier stage than differential expression significance. Significance thresholds, effect size parameters, and multiple testing corrections are described under each sub-analysis above.

3. RESULTS

3.1 Multi-Tissue Chromatin Accessibility Atlas of Human Ontogenesis

Following quality filtering, a total of 412 847 high-confidence ATAC-seq peaks were identified across all 75 specimens. The median fraction of reads in peaks (FRiP) score was 0.48 (IQR 0.42–0.55), indicating high library quality. On UMAP embedding, specimens segregated primarily by tissue lineage (explaining 61.3% of variance in the first two UMAP dimensions) and secondarily by ontogenetic stage (27.8%), confirming that tissue identity is a stronger determinant of chromatin architecture than developmental stage throughout the studied window.

Peak classification by specificity revealed that 50.4% of peaks were constitutively accessible across all tissue-stage combinations, likely corresponding to housekeeping gene promoters and broadly active regulatory elements. Of the remaining 49.6% conditionally accessible peaks, 38.4% were stage-specific (differentially accessible in at least one stage comparison regardless of tissue) and 11.2% were tissue-specific (differentially accessible in at least one tissue comparison regardless of stage). A small subset of peaks (3.7%) displayed both strong stage and tissue specificity, representing the most developmentally informative regulatory elements. These jointly stage-tissue-specific peaks were highly enriched at known developmental gene loci including GATA4, NKX2-5, PAX6, FOXA2, and MYOD1 (adjusted $p < 0.001$ for all; hypergeometric test).

3.2 Chromatin Accessibility Precedes Transcriptional Activation

To quantify the temporal relationship between chromatin opening and gene activation, we paired each differentially accessible peak with its nearest differentially expressed gene across

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

consecutive stage transitions. In 71.3% of loci achieving differential expression between any two stages, the cognate regulatory peak was already differentially accessible in the preceding ontogenetic stage, confirming that chromatin remodelling anticipates transcriptional output by at least one developmental stage. This 'accessibility-first' pattern was most pronounced in the neural (79.6% of loci) and hepatic (74.1%) lineages and least pronounced in skeletal muscle (62.8%), where transcriptional activation appeared to co-occur with chromatin opening more frequently, potentially reflecting the more abrupt nature of myogenic commitment driven by the MyoD transcriptional cascade [6].

3.3 Lineage-Specific Transcription Factor Networks

SCENIC+ analysis identified 1 847 eRegulons across all tissue-stage combinations. The 20 highest-RSS eRegulons per tissue at each stage are listed in Supplementary Table S1; key findings are summarised in Table 1. In the hepatic lineage, SOX17 and FOXA2 ('pioneer factors') dominated Stage I accessibility, opening enhancers associated with albumin, transferrin, and cytochrome P450 loci; their regulon activity declined at Stage III as more terminal hepatocyte identity factors (HNF4A, CEBPA) ascended. In the cardiac lineage, NKX2-5 and GATA4 formed a co-regulatory module maximally active at Stage II, consistent with their established roles in cardiomyocyte specification and trabeculation. The neural cortical lineage displayed sequential TF dominance: PAX6 (Stage I progenitor maintenance), NEUROG2 (Stage II intermediate progenitor), and NEUROD1 (Stage III committed neuron), recapitulating the canonical cortical neurogenesis cascade. Renal lineage progression was characterised by WT1, PAX2, and FOXD1, consistent with nephron progenitor identity and commitment. Skeletal muscle commitment was dominated by MYOD1 and MYOGENIN across Stages II and III, with MEF2C serving as a co-activator at Stage III.

Table 1. Master transcription factor regulons by tissue lineage and ontogenetic stage (top 3 per tissue-stage combination by regulon specificity score).

Tissue Lineage	Stage	Top TF Regulon 1	Top TF Regulon 2	Top TF Regulon 3
Hepatic	I (ED 28–42)	SOX17 (RSS 0.91)	FOXA2 (RSS 0.88)	GATA6 (RSS 0.82)
Hepatic	II (ED 56–70)	HNF4A (RSS 0.87)	FOXA1 (RSS 0.84)	CEBPB (RSS 0.79)
Hepatic	III (FW 16–20)	HNF4A (RSS 0.93)	CEBPA (RSS 0.90)	NR1H4 (RSS 0.81)
Neural cortical	I (ED 28–42)	PAX6 (RSS 0.93)	SOX2 (RSS 0.90)	HES1 (RSS 0.85)
Neural cortical	II (ED 56–70)	NEUROG2 (RSS 0.91)	EOMES (RSS 0.87)	PAX6 (RSS 0.74)

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

Neural cortical	III (FW 16–20)	NEUROD1 (RSS 0.94)	TBR1 (RSS 0.88)	MEF2C (RSS 0.80)
Cardiac	I (ED 28–42)	GATA4 (RSS 0.89)	TBX5 (RSS 0.84)	MEF2A (RSS 0.79)
Cardiac	II (ED 56–70)	NKX2-5 (RSS 0.95)	GATA4 (RSS 0.91)	TBX20 (RSS 0.85)
Skeletal muscle	II (ED 56–70)	MYOD1 (RSS 0.96)	MYF5 (RSS 0.89)	MYOG (RSS 0.82)
Renal	I (ED 28–42)	WT1 (RSS 0.90)	PAX2 (RSS 0.87)	FOXD1 (RSS 0.83)

RSS: regulon specificity score (0–1 scale; higher values indicate greater lineage and stage specificity). TF: transcription factor; ED: embryonic day; FW: fetal week. Only tissue-stages with the highest RSS scores are shown; full regulon tables are available in Supplementary Data.

3.4 Bivalent Domain Resolution During Ontogenetic Progression

To assess the fate of developmentally poised chromatin during ontogenesis, we integrated ATAC-seq accessibility data with paired H3K4me3 and H3K27me3 ChIP-seq profiles from the subset of 30 specimens. At Stage I, 14 283 gene promoters were classified as bivalent (H3K4me3+/H3K27me3+ co-marked with low ATAC-seq accessibility). Between Stage I and Stage III, 68.4% of these bivalent domains were resolved: 41.2% transitioned to a fully active state (H3K4me3-only, high ATAC accessibility, elevated RNA expression) and 27.2% transitioned to a stably repressed state (H3K27me3-only, low ATAC accessibility, absent RNA expression). The remaining 31.6% retained bivalency at Stage III, predominantly at genes encoding transcription factors required for non-adopted lineages, consistent with a cellular memory function that preserves latent plasticity [4]. Lineage-inappropriate fate gene promoters (e.g., NKX2-5 in neural cells; PAX6 in cardiac cells) showed the highest rates of bivalency retention (82.7%) across all tissues.

3.5 Super-Enhancer Emergence and Lineage Identity

Super-enhancer (SE) calling identified 2 416 SEs across all tissue-stage combinations, with a median of 163 SEs per tissue at Stage III (range 141–184). The number of SEs per tissue increased significantly from Stage I to Stage III (median 89 at Stage I vs. 163 at Stage III; paired t-test $p < 0.001$), consistent with progressive consolidation of lineage identity during ontogenetic commitment. The most Stage-III-specific SEs were associated with cardinal identity genes of each lineage: ALB and CYP3A4 in hepatic tissue; SCN5A and MYH7 in cardiac; NEFM and SYP in neural; MYH2 and TNNT3 in skeletal muscle; and UMOD and SLC12A1 in renal tissue. Strikingly, 74.1% of Stage-III SEs overlapped with ATAC-seq peaks that were already accessible at Stage II, confirming that SE establishment follows from prior chromatin opening rather than preceding it, and further validating the temporal accessibility-first model of ontogenetic regulation. Key quantitative metrics across tissues and stages are summarised in Table 2.

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

Table 2. Chromatin accessibility and transcriptome metrics across tissue lineages and ontogenetic stages.

Tissue	Stage	ATAC Peaks (×10³)	Stage-Specific Peaks (%)	DEGs (vs. Stage I)	Super-Enhancers (n)	Bivalent Loci (%)
Hepatic	I	68.4	—	—	91	38.4
Hepatic	II	74.2	21.3	2 847	128	24.7
Hepatic	III	81.7	34.8	5 614	161	12.1
Neural cortical	I	71.2	—	—	84	41.2
Neural cortical	III	88.3	41.2	6 203	184	10.8
Cardiac	I	64.8	—	—	79	36.9
Cardiac	III	76.4	36.1	4 988	158	13.4
Skeletal muscle	III	69.3	29.7	4 312	141	15.9
Renal	III	72.1	31.5	4 741	153	14.2

DEGs: differentially expressed genes relative to Stage I. Bivalent loci (%): percentage of Stage I bivalent (H3K4me3+/H3K27me3+) domains retaining bivalency at the indicated stage. Stage II data shown only for hepatic lineage in full; other lineages shown at Stage I and Stage III for brevity.

4. DISCUSSION

This study presents the first integrated ATAC-seq and RNA-seq atlas of human somatic tissue ontogenesis spanning five lineages and three defined developmental stages, generated within a unified analytical framework. The principal findings—that chromatin accessibility temporally precedes transcriptional activation in the majority of loci, that lineage-specific TF networks exhibit precise stage-dependent dominance hierarchies, that bivalent domains resolve progressively into active or repressed states, and that super-enhancers accumulate and consolidate during ontogenetic commitment—collectively advance a hierarchical model of

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

human developmental epigenomics with important implications for regenerative medicine and developmental biology.

The finding that 71.3% of gene loci showed prior chromatin opening before transcriptional activation is consistent with the concept of 'epigenetic priming' articulated by Reik and colleagues [3], who proposed that the sequential deployment of pioneer transcription factors—capable of binding nucleosomal DNA and displacing or repositioning nucleosomes—creates permissive chromatin landscapes that enable subsequent transcriptional regulators to activate gene expression. In our dataset, the hepatic and neural lineages showed the strongest accessibility-first patterns, which is coherent with the known biology of hepatic pioneer factors FOXA2 and SOX17, and neural cortical pioneer factor PAX6, each of which has been independently shown to remodel chromatin in a ligand-independent, DNA sequence-driven manner prior to the recruitment of downstream activating complexes [1].

The near-perfect sequential TF dominance hierarchy in the neural cortical lineage—PAX6 (Stage I) → NEUROG2 (Stage II) → NEUROD1 (Stage III)—recapitulates at the chromatin accessibility level the canonical transcriptional cascade of human cortical neurogenesis established from single-cell RNA-seq studies of the developing cortex [7]. This convergence provides strong cross-validation for the regulatory inference performed by SCENIC+, confirming that eRegulon activity scores derived from the joint ATAC-seq/RNA-seq signal accurately reflect biologically meaningful TF activity states in vivo. Importantly, our data extend this cascade by demonstrating that NEUROG2 eRegulon activation is associated with the opening of distal enhancers (median distance from transcription start site 34.2 kb) rather than promoter-proximal elements, emphasising the dominant role of long-range regulatory interactions in neural commitment.

The resolution of bivalent chromatin domains during ontogenetic progression warrants particular discussion in the context of developmental plasticity and cellular reprogramming. The 31.6% of domains retaining bivalency at Stage III—preferentially at lineage-inappropriate fate gene promoters—suggests that fetal somatic cells are not epigenetically fully locked into their committed identity and retain some degree of latent reprogramming potential. This finding is consistent with the greater ease of generating induced pluripotent stem cells (iPSCs) from fetal compared with adult somatic cells [4] and may reflect the incomplete deposition of constitutive heterochromatin (H3K9me3, H3K9me2) at non-adopted fate loci during the developmental window studied here. Whether these bivalent domains are fully resolved to stable H3K9me3-mediated silencing in postnatal tissues, and whether their retention in fetal cells represents a targetable feature for regenerative reprogramming, are important questions for future investigation.

The progressive accumulation of super-enhancers from Stage I to Stage III, and the finding that 74.1% of Stage-III SEs were preceded by ATAC-seq accessibility at Stage II, provides mechanistic insight into how SE establishment occurs during ontogenesis. Rather than de novo chromatin opening at Stage III, SEs appear to consolidate at regulatory regions that were already accessible—but perhaps not yet fully activated—at the preceding stage. This is consistent with a model in which the arrival of high-abundance master TFs (such as NKX2-5 in cardiac or MYOD1 in skeletal muscle) converts pre-accessible enhancer regions into fully activated, SE-level transcriptional hubs by driving phase-separated condensate formation around densely bound enhancer clusters [5]. The specific molecular events governing this accessibility-to-SE transition—including the role of BRD4, MED1, and cohesin in condensate nucleation at pre-accessible developmental enhancers—represent a compelling mechanistic frontier.

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

Several limitations of this study should be acknowledged. First, the use of bulk tissue specimens rather than single-cell resolution methods means that cell-type heterogeneity within each tissue introduces noise into the chromatin accessibility and gene expression profiles. Particularly at Stage I and II, when tissue cellular composition is dynamically changing and progenitor cells coexist with early post-mitotic derivatives, bulk profiling averages over distinct cellular states. Future studies using single-cell ATAC-seq (scATAC-seq) and single-cell RNA-seq (scRNA-seq) will be needed to resolve cell-type-specific regulatory dynamics within these stages. Second, the ChIP-seq data for histone modification analysis was available only for a subset of specimens, limiting the statistical power of bivalency and SE analyses. Third, the three ontogenetic stages studied, while capturing important developmental transitions, do not encompass the neonatal and postnatal maturation stages during which additional epigenetic consolidation occurs [8]. Fourth, functional validation of the inferred eRegulon activities—through TF perturbation experiments in matched organoid systems—was beyond the scope of the present study.

5. CONCLUSION

This study establishes the first integrated multi-tissue epigenomic atlas of human somatic ontogenesis and demonstrates that tissue-specific developmental gene programmes are governed by a hierarchical regulatory logic in which: (i) chromatin accessibility established by pioneer transcription factors temporally precedes and predicts transcriptional activation; (ii) lineage-specific TF networks exhibit precise stage-dependent dominance cascades that mirror and extend established developmental paradigms; (iii) bivalent chromatin domains resolve progressively toward active or repressed states, with lineage-inappropriate fate gene promoters preferentially retaining bivalency; and (iv) super-enhancers consolidate during commitment at pre-accessible regulatory regions established in earlier ontogenetic stages.

Together, these findings provide a mechanistic framework for understanding how the shared genome of all somatic cells is differentially interpreted to generate the tissue diversity that characterises the adult human organism. The atlas generated here—comprising 412 847 annotated chromatin accessibility peaks, transcriptome profiles across 75 specimens, and 1 847 TF eRegulons—constitutes a high-resolution reference resource for the developmental biology community. Practically, this resource will facilitate the optimisation of in vitro differentiation protocols for regenerative medicine by identifying the TF and epigenetic checkpoints that must be traversed in the correct sequence to recapitulate authentic ontogenetic trajectories in stem cell-derived organoids.

Future research priorities emerging from this study include: (i) single-cell resolution re-analysis of the identified developmental transitions to resolve cell-type-specific regulatory heterogeneity; (ii) extension of profiling to neonatal and early postnatal stages to complete the prenatal-to-postnatal ontogenetic continuum; (iii) functional validation of identified master TF regulons by CRISPR interference and overexpression in lineage-matched organoid systems; (iv) integration of 3D chromatin organisation data (Hi-C, Micro-C) to characterise the topologically associating domain dynamics that enable long-range SE-to-promoter communication during ontogenetic progression; and (v) cross-species comparative analysis to identify evolutionarily conserved versus human-specific epigenetic regulatory mechanisms in somatic tissue ontogenesis.

REFERENCES

pISSN:2349-5707 Volume:13,Issue 5, May -2026 eISSN:2349-5715

1. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011;25(21):2227–2241. doi:10.1101/gad.176826.111.

2. Corces MR, Trevino AE, Hamilton EG, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods. 2017;14(10):959–962. doi:10.1038/nmeth.4396.

3. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature. 2007;447(7143):425–432. doi:10.1038/nature05918.

4. Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. Genes Dev. 2013;27(12):1318–1338. doi:10.1101/gad.219626.113.

5. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307–319. doi:10.1016/j.cell.2013.03.035.

6. Tapscott SJ. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. Development. 2005;132(12):2685–2695. doi:10.1242/dev.01874.

7. Nowakowski TJ, Bhaduri A, Pollen AA, et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science. 2017;358(6368):1318–1323. doi:10.1126/science.aap8809.

8. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–330. doi:10.1038/nature14248.