



Epigenomic states contribute to coordinated allelic transcriptional bursting in iPSC reprogramming

Parichitran Ayyamperumal, Hemant Chandru Naik, Amlan Jyoti Naskar, Lakshmi Sowjanya Bammidi, Srimonta Gayen

Two alleles of a gene can be transcribed independently or coordinately, which can lead to temporal expression heterogeneity with potentially distinct impacts on cell fate. Here, we profiled genome-wide allelic transcriptional burst kinetics during the reprogramming of MEF to induced pluripotent stem cells. We show that the degree of coordination of allelic bursting differs among genes, and alleles of many reprogramming-related genes burst in a highly coordinated fashion. Notably, we show that the chromatin accessibility of the two alleles of highly coordinated genes is similar, unlike the semi-coordinated or independent genes, suggesting the degree of coordination of allelic bursting is linked to allelic chromatin accessibility. Consistently, we show that many transcription factors have differential binding affinity between alleles of semi-coordinated or independent genes. We show that highly coordinated genes are enriched with chromatin accessibility regulators such as H3K4me3, H3K4me1, H3K36me3, H3K27ac, histone variant H3.3, and BRD4. Finally, we demonstrate that enhancer elements are highly enriched in highly coordinated genes. Our study demonstrates that epigenomic states contribute to coordinated allelic bursting to fine-tune gene expression during induced pluripotent stem cell reprogramming.

DOI [10.26508/lisa.202302337](https://doi.org/10.26508/lisa.202302337) | Received 24 August 2023 | Revised 30 January 2024 | Accepted 30 January 2024 | Published online 6 February 2024

Introduction

Reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) serves as an excellent model system for studying the mechanisms of cell fate specification and gene regulation (Takahashi & Yamanaka, 2006, 2016; Onder et al, 2012; Polo et al, 2012; Cheloufi et al, 2015; Bauer et al, 2021; Generoso et al, 2023; Naik et al, 2023 *Preprint*). However, mechanisms of iPSC reprogramming remain poorly understood. Precise expression of genes involved in iPSC reprogramming is crucial for reprogramming. To date, the gene expression dynamics during iPSC reprogramming have been studied at the gene level but not at the allelic level. In eukaryotic cells, transcription happens in a sporadic manner through random transcriptional bursting separated by periods of silent state, which contributes to the gene expression

heterogeneity among the identical cells (Raj et al, 2006; Raj & van Oudenaarden, 2008; Little et al, 2013; Padovan-Merhar et al, 2015; Larsson et al, 2019b; Naik et al, 2021; RV et al, 2021). On the other hand, the regulation of transcriptional bursting can be shared or autonomous between alleles (Onuchic et al, 2018; Finn et al, 2019; Naik et al, 2021; RV et al, 2021). Indeed, recent studies by us and others demonstrated that often the kinetics of transcriptional bursting of the two alleles of a gene differs, which in turn gives rise to dynamic autosomal random monoallelic expression (aRME) (Borel et al, 2015; Reinius & Sandberg, 2015; Gendrel et al, 2016; Reinius et al, 2016; Gregg, 2017; Naik et al, 2021; RV et al, 2021). It is believed that dynamic aRME can induce temporal variations of gene expression among cells and thereby may contribute to cell fate specification. Therefore, it is important to profile the transcriptional kinetics of genes at the allelic level to understand how gene expression is fine-tuned for precise cell fate specification. However, allelic transcriptional kinetics during iPSC reprogramming remains unexplored. To address this, we have profiled genome-wide transcriptional burst kinetics at allelic level across different stages of reprogramming of MEF to iPSC using allele-specific single-cell RNA-sequencing (scRNA-seq) analysis. We find that the degree of coordination of allelic bursting differs among genes and changes dynamically during iPSC reprogramming. Importantly, we find that many genes involved in reprogramming pathways have a high degree of allelic coordination. On the other hand, the factors involved in coordinating allelic bursting are not known. Transcription factors and enhancer functions are thought to modulate transcriptional burst kinetics (Larsson et al, 2019b). Emerging trends suggest that epigenomic states are linked to burst kinetics (Nicolas et al, 2018; Chen et al, 2019; Fraser et al, 2021). Therefore, we have explored how epigenomic features contribute to the coordinated allelic transcriptional bursting during iPSC reprogramming.

Result

Prevalent bursty expression contributes to dynamic aRME during iPSC reprogramming

Transcription of many genes occurs in a stochastic manner, where genes undergo sporadic bursting to produce RNA. The kinetics of

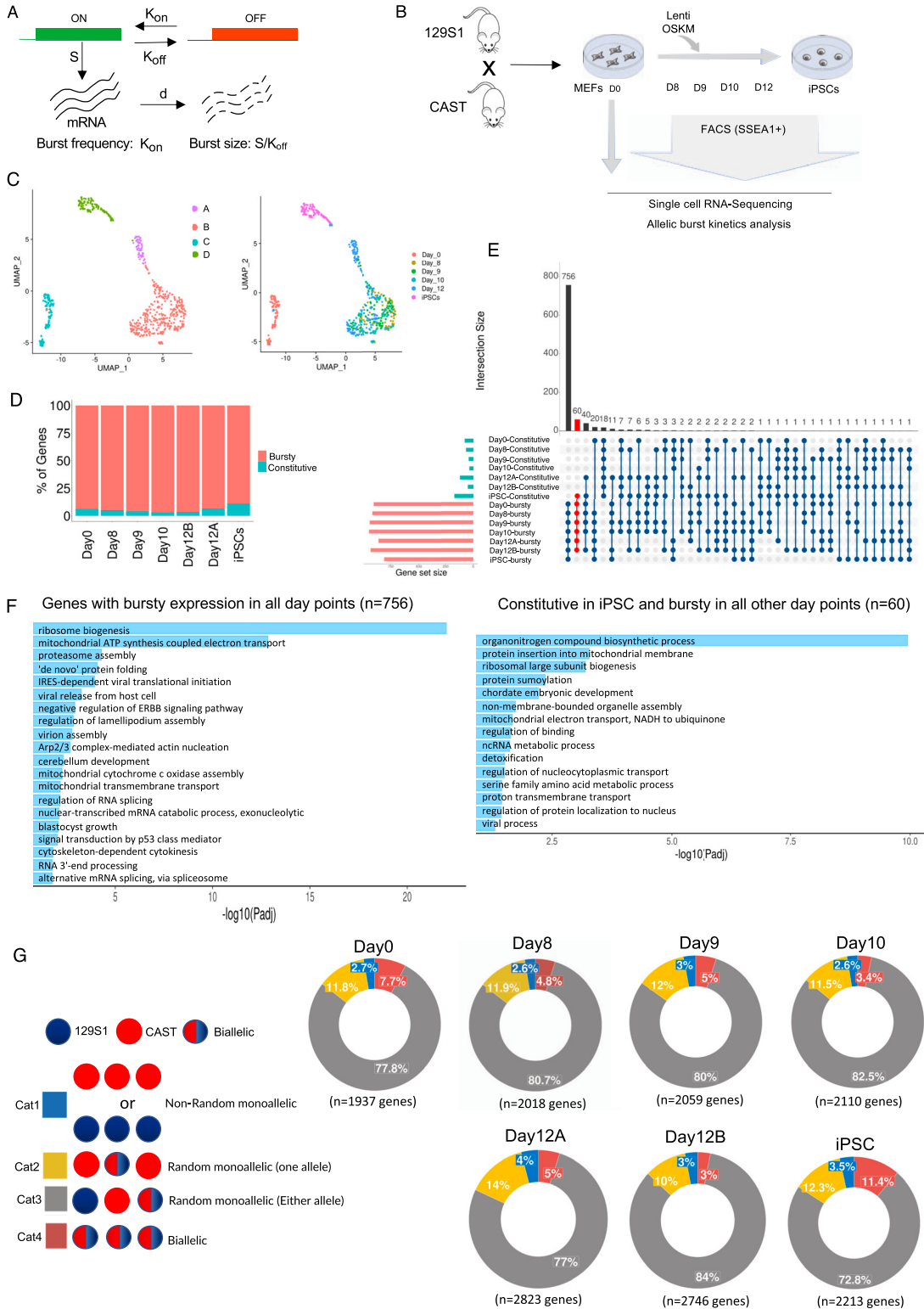


Figure 1. Genome-wide profiling of bursty expression and dynamic autosomal random monoallelic expression in different stages of MEF to induced pluripotent stem cell (iPSC) reprogramming.

(A) A diagrammatical representation of the “two-state model” of transcriptional bursting. K_{on} is the rate of transcriptional activation of a gene; K_{off} is the rate at which a gene becomes transcriptionally inactive; S is the rate of transcription of a gene in an active state; d is the rate of mRNA decay; burst kinetics is represented through K_{on} (burst frequency) and S/K_{off} (burst size). (B) Graphical representation of OSKM (Oct-3/4, Sox2, Klf4, and c-Myc)-mediated reprogramming of hybrid MEF cells (129S1 X CAST) and collection of cells of intermediate stages day 8, day 9, day 10, day 12, and iPSCs. (C) Plots representing UMAP-based clustering of cells of the different stages of

transcriptional bursting are deduced through the well-known “two-state model” of transcription. According to the “two-state model,” the promoter of a gene switches stochastically from an inactive/OFF state to an active/ON state, and burst kinetics is determined through two parameters: burst frequency and burst size (Fig 1A) (Chubb et al, 2006; Raj & van Oudenaarden, 2008; Larson, 2011). Burst frequency is defined by the number of bursts per unit time, and burst size explains the average amount of mRNA molecules per burst (Fig 1A). To date, burst kinetics have been studied mainly at the gene level, not the allelic level. However, profiling burst kinetics at allelic resolution is important as the kinetics of bursting between alleles often differ (Naik et al, 2021). Therefore, to understand how individual alleles of a gene contribute to the overall gene expression, we have delineated the kinetics of allelic bursting genome-wide across different stages of reprogramming of MEF to iPSC (Fig 1B). To profile, genome-wide allelic transcriptional burst kinetics, we performed SCALE (Single-Cell ALlelic Expression) analysis using scRNA-seq datasets (Fig 1B). SCALE relies on the Empirical Bayes Framework, which first classifies the genes into monoallelic, biallelic, and silent based on the allele-specific read counts across the cells. Subsequently, the biallelic genes are further categorized into biallelic bursty and biallelic constitutive (Fig S1A) (Jiang et al, 2017). Biallelic genes that have zero expression of both alleles in a few cells but not all cells are considered bursty genes, whereas genes that have at least one allele constitutive expression or both allele constitutive expression across cells are considered biallelic constitutive genes (Fig S1A). The MEF cells used in the experiment were derived from a cross of two divergent mouse strains *M. Musculus* (129S1) and *M. Castaneus* (CAST), thereby enabling us to perform allele-specific analysis based on strain-specific SNPs (Fig 1B) (Talon et al, 2021). We excluded low-expressed genes from our study to avoid allelic dropout-related technical noise, which could lead to inaccurate estimation of the allelic expression (Kim et al, 2015; Santoni et al, 2017; Zhao et al, 2017).

First, we performed UMAP-based clustering to identify homogeneous cell populations across the different stages of reprogramming. We found that day 0-MEF cells and iPSC cells formed individual tight clusters (Fig 1C). Intermediate cells (day 8, day 9, day 10) also clustered together. However, day 12 cells formed two different sub-clusters (day 12A and day 12B). Therefore, we performed SCALE analysis separately for day 12A and day 12B sub-clusters to reduce the effect of heterogeneity in our burst kinetics analysis (Fig 1C). Through SCALE analysis, we found that most of the biallelic genes (~90%) have bursty expression across all stages of reprogramming (Fig 1D). Next, we compared the bursty/constitutive pattern of gene expression across different reprogramming stages and found that most genes ($n = 756$) maintained bursty expression throughout the reprogramming (Fig 1E). Interestingly, these genes were enriched toward key reprogramming-related processes such as ribosome biogenesis, proteasome assembly, blastocyst growth,

etc. (Fig 1F). Moreover, genes ($n = 60$) that remained bursty on all days except iPSCs were also enriched in many reprogramming-related biological processes, including chordate embryonic development (Fig 1F). These results indicated that many genes involved in iPSC reprogramming pathways exhibit bursty expression. However, a few genes ($n = 20$) that maintained bursty expression in all-day points except MEF did not show enrichment towards major iPSC reprogramming pathways (Fig S1B). Genes ($n = 18$), which remained constitutive throughout the reprogramming, were not enriched to major reprogramming pathways (Fig S1B). On the other hand, we have previously shown that bursty expression results in dynamic aRME, which creates gene expression heterogeneity among cells (Naik et al, 2021). Therefore, we explored the landscape of allelic expression throughout the different stages of iPSC reprogramming by allele-specific scRNA-seq analysis. We categorized a gene as monoallelic within a cell if at least 95% of the expression originated from one allele. We classified genes based on their allelic expression patterns into four categories: Cat 1: nonrandom monoallelic (same allele expressed across the cells); Cat 2: RME with one allele (genes expressed biallelically in some cells whereas exhibiting monoallelic expression from one allele consistently in other cells); Cat 3: random monoallelic with either allele (genes expressed biallelically in some cells whereas exhibiting monoallelic expression from either of the allele in other cells); and Cat 4: biallelic (genes expressed biallelically across the cells) (Fig 1G). Our analysis revealed that ~80–90% of genes have dynamic aRME across different stages of reprogramming (Fig 1G). Analysis of allelic expression of X-linked genes in MEF showed monoallelic expression from the CAST allele as 129S1-X is inactivated in these cells, thereby validating our allele-specific analysis pipeline (Fig S1C). X-inactivation is a process through which female mammals compensate for the dosage of X-linked gene expression between sexes (Gayen et al, 2015, 2016; Sarkar et al, 2015; Saiba et al, 2018; Kaur et al, 2020; Samanta et al, 2022). Altogether, our analysis suggested widespread bursty expression of genes during iPSC reprogramming, resulting in dynamic aRME.

Alleles of genes exhibit similar burst kinetics but have different coordination during iPSC reprogramming

Next, we investigated if the two alleles of a gene have similar burst kinetics or not. To explore this, we profiled burst frequency and burst size at the allelic level for the biallelic bursty genes using SCALE. SCALE relies on a “two-state model” of transcription where a gene switches from active to inactive state with activation rate of Kon and deactivation rate of $Koff$. When a gene is in the active state, the rate of transcription is S and the rate of RNA decay is d (Fig 1A). Burst frequency is determined by the number of bursts per unit time (Kon), and burst size is deduced through average number of mRNA molecules per burst when the gene is in active state ($S/Koff$)

reprogramming. (D) Quantification of the percent of genes exhibiting bursty expression across all-day points of reprogramming. The number of bursty genes in day 0: 1,769 (of 1,886 biallelic genes), day 8: 1,869 (of 1,968 biallelic genes), day 9: 1,921 (of 2,004 biallelic genes), day 10: 1,993 (of 2,058 biallelic genes), day 12A: 2,544 (of 2,723 biallelic genes), day 12B: 2,585 (of 2,678 biallelic genes), and iPSC: 1,901 (of 2,137 biallelic genes). (E) Cross-comparison plot of bursty and constitutive genes across the different stages of reprogramming. (F) Gene ontology (GO) enrichment analysis of 756 genes that remain bursty across all-day points and 60 genes that are constitutive in iPSCs but bursty in other-day points. (G) Plots representing the percent of genes with different allelic expression categories throughout different stages of reprogramming; Cat 1: nonrandom monoallelic, Cat 2: random monoallelic with one allele, Cat 3: random monoallelic with either allele, Cat 4: biallelic.

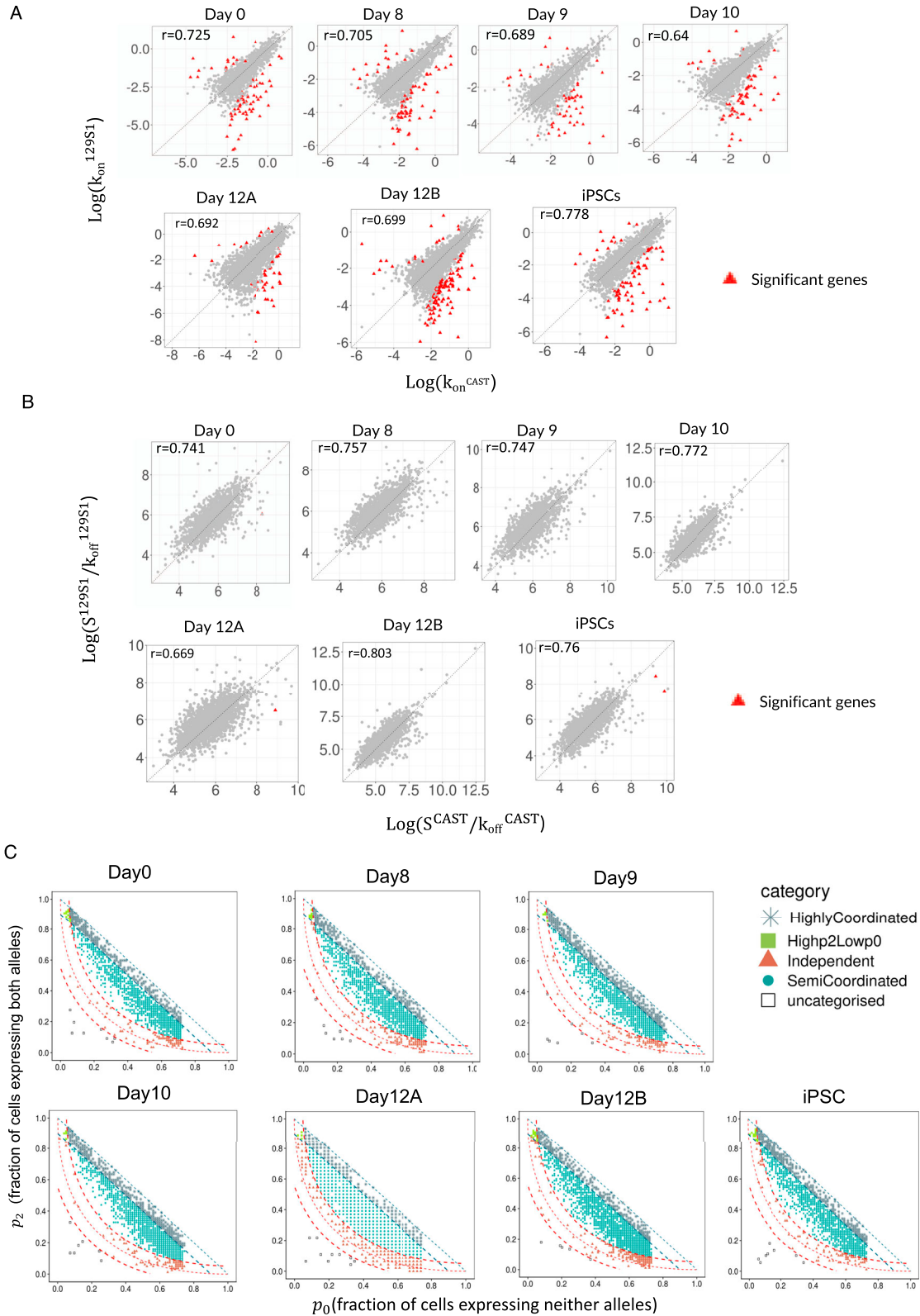


Figure 2. Profiling of allelic burst kinetics and coordination.

(A, B) Plots representing correlation between (A) allelic burst frequency across different stages of reprogramming day 0 MEF ($r = 0.725$), day 8 ($r = 0.705$), day 9 ($r = 0.689$), day 10 ($r = 0.64$), day 12A ($r = 0.69$), day 12B ($r = 0.69$) and induced pluripotent stem cells (iPSCs) ($r = 0.778$) and (B) allelic burst size in day 0 MEF ($r = 0.741$), day 8 ($r = 0.757$), day 9 ($r = 0.747$), day 10 ($r = 0.772$), day 12A ($r = 0.66$), day 12B ($r = 0.80$) and iPSCs ($r = 0.76$). Genes that exhibit significant differences in burst frequency and size between two

(Fig 1A). We observed a high degree of correlation of both burst frequency ($r = 0.64\text{--}0.778$) and burst size ($r = 0.669\text{--}0.803$) between two alleles across all-day points (Fig 2A and B). Very few genes exhibited significant burst frequency and size differences between the two alleles, as marked by the red triangles (Fig 2A and B). Taken together, our results suggest that alleles of most of the genes have similar burst kinetics. Next, we explored the degree of coordination of bursting between two alleles by plotting the percent of cells expressing neither allele (p_0) versus the percent of cells expressing both alleles (p_2) (Fig 2C). Blue diagonal line represents perfect coordination ($p_0 + p_2 = 1$), whereas the red curve signifies independent bursting with shared kinetics (Fig 2C). We categorized genes based on the degree of coordination of allelic bursting into three categories: (1) highly coordinated genes: with $p_0 + p_2 > 0.90$ marked by gray asterisk between blue dotted diagonal lines; (2) independent genes: genes near the red curve with a threshold of $+0.05$, signified by the upper red curve, and -0.05 , signified by the lower red curve, marked by rosewood triangles; (3) semi-coordinated genes that lie between the uppermost red curve line and lower blue dotted diagonal line (Persian blue dots) (Fig 2C).

We observed that most of the genes exhibited semi-coordinated allelic bursting at all-day points. However, many genes also showed highly coordinated allelic bursting across the different stages of reprogramming (Fig 2C).

Genes involved in iPSC reprogramming undergo highly coordinated allelic bursting

Next, we performed a gene ontology (GO) enrichment analysis of highly coordinated genes across all stages of reprogramming (Fig 3). Interestingly, we found that highly coordinated genes in iPSCs are enriched in processes linked to stem cell population maintenance, in-utero embryo development, endoderm formation, etc. (Fig 3A, Supplemental Data 1). Furthermore, highly coordinated genes in iPSCs are enriched in cellular respiration, cristae formation, and glutamine metabolism, which are relevant to metabolic remodeling in iPSCs (Fig 3A) (Teslaa & Teitell, 2015; Mathieu & Ruohola-Baker, 2017; Tohyama et al, 2017). Importantly, genes involved in ribosome biogenesis, crucial for stem cell maintenance, were highly enriched into the highly coordinated gene cohort in iPSC (Fig 3A) (Gabut et al, 2020). On the contrary, highly coordinated genes in day 0 MEF cells were not enriched in such related pathways. Interestingly, across the different intermediate stages of reprogramming, highly coordinated genes were enriched towards many reprogramming-related biological processes such as ribosome biogenesis, protein folding, stem cell maintenance, NFkB signaling, etc. (Fig 3B) (Gabut et al, 2020; Yan et al, 2020; Kaltschmidt et al, 2021). Next, we performed a cross-comparison of the three allelic coordination categories across all-day points and found that,

whereas many genes maintained a similar degree of allelic coordination across all-day points, some genes did not (Fig S2A). Strikingly, we found that the genes ($n = 52$) that become highly coordinated in iPSCs are enriched towards ribosome biogenesis, aerobic respiration, nuclear pore formation, which are highly critical and directly relevant for iPSC reprogramming (Fig 3C) (Xu et al, 2013; Teslaa & Teitell, 2015). Moreover, genes that remain highly coordinated ($n = 25$) on all days showed enrichment towards some iPSC reprogramming-linked functions like protein stability, NFkB pathway, mitochondrion organization, etc. (Fig 3C). Altogether, our analysis revealed that the two alleles of many genes involved in iPSC reprogramming have a high degree of transcriptional bursting coordination (Supplemental Data 1). In parallel, we also observed that genes that remained semi-coordinated ($n = 266$) on all days or converted from highly coordinated on day 0 to semi-coordinated on other days ($n = 33$) were enriched towards some iPSC reprogramming-related pathways like proteasome assembly, oxidative stress response, etc. (Fig S2B) (Szutorisz et al, 2006; Schröter & Adjaye, 2014; Hawkins et al, 2016; Mathieu & Ruohola-Baker, 2017). Next, we investigated if higher coordination of transcriptional bursting between alleles of individual genes facilitates the optimal expression dosage of these highly coordinated genes involved in iPSC reprogramming. To explore this, we compared the overall expression level of highly coordinated, semi-coordinated, and independent genes. We found that the indeed highly coordinated genes have higher expression level compared to the semi-coordinated and independent genes (Fig 3D). Taken together, we conclude that highly coordinated allelic bursting helps to fine-tune the optimal expression of genes involved in iPSC reprogramming.

Coordinated allelic bursting is linked to chromatin accessibility

To understand the mechanisms of allelic bursting coordination, we asked whether the degree of coordination of allelic bursting is linked to epigenomic states. To address this, we profiled genome-wide allelic chromatin accessibility across different stages of MEF to iPSC reprogramming through allele-specific analysis of available ATAC-seq (ATAC-seq) datasets (Talon et al, 2021). The same hybrid MEFs (129S1X CAST), as described for scRNA-seq, were used for this experiment, allowing us to profile chromatin accessibility at the allelic level. We analyzed ATAC-seq in MEFs (day 0) and across reprogramming stages (SSEA1+ reprogramming intermediates at days 8, 9, 10, 12, and iPSCs), like the burst kinetics analysis (Fig 4A). We first validated our allele-specific ATAC-seq analysis pipeline by quantifying the difference in the enrichment of ATAC-seq reads between active-X (129S1) versus inactive-X (CAST allele) (Fig S3A). In consistence with previous reports, in MEFs and early reprogramming intermediates, active-X (CAST) showed strong enrichment of ATAC-seq reads, whereas the inactive-X (129S1) showed almost no

alleles have been marked by red triangles. (C) Smooth scatterplots representing bursting coordination between two alleles of genes for day 0 MEF, day 8, day 9, day 10, day 12A, day 12B, and iPSCs. Percent of cells expressing neither allele (p_0) is plotted with the percent of cells expressing both alleles (p_2); the blue diagonal line represents perfect coordination ($p_0 + p_2 = 1$), whereas the red curve signifies independent bursting with shared kinetics. Different categories of genes based on allelic bursting coordination: low p_0 high p_2 (green filled squares), perfectly coordinated ($p_0 + p_2 > 0.90$ marked by gray asterisk between blue dotted diagonal lines), independent genes marked by rosewood triangles (between upper and lower red curved lines, with a threshold of $+0.05$ signified by upper red curve and -0.05 , signified by the lower red curve), and semi-coordinated genes marked with persian blue dots.

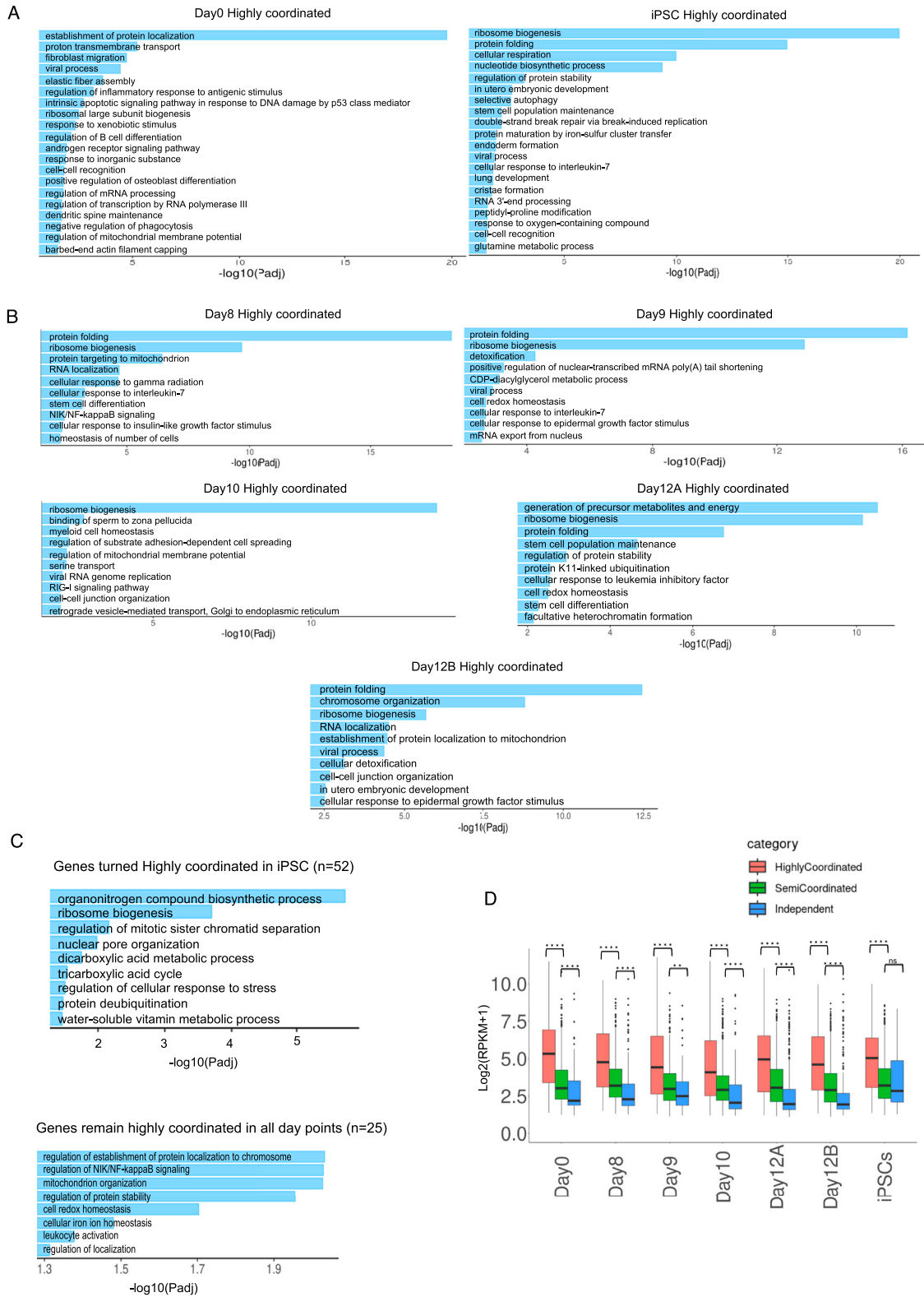


Figure 3. Highly coordinated genes are enriched to induced pluripotent stem cell (iPSC) reprogramming-related processes. (A) Gene ontology (GO) enrichment analysis of highly coordinated genes in day 0 MEF and iPSCs. (B) Gene ontology (GO) enrichment analysis of highly coordinated genes in intermediate stages of reprogramming: day 8, day 9, day 10, and day 12 cells. (C) Gene ontology (GO) enrichment analysis of genes (n = 52) that become highly coordinated in iPSCs (top) and genes (n = 25) that maintain highly coordinated allelic bursting through all-day points. (D) Plot representing the expression level (reads per

enrichment, and upon reactivation of the inactive-X towards the attainment of iPSCs, there was a gain of chromatin accessibility (Fig S3A). Taken together, enrichment analysis of ATAC-seq reads of X-linked genes validated the accuracy of our method. Next, we compared the enrichment of ATAC-seq reads between two alleles of different categories of genes (highly coordinated, semi-coordinated, and independent) across the gene body and 3 kb upstream of TSS and 3 kb downstream of TES during reprogramming (Fig 4A). Interestingly, our analysis revealed that the two alleles of highly coordinated genes have very similar enrichment at most day points. Whereas enrichment of ATAC-seq reads of alleles of semi-coordinated/independent genes differed in most cases (Fig 4A). As expected, allelic enrichment of ATAC-seq reads considering all autosomal genes was quite similar (Fig 4A).

Notably, genes that maintained semi-coordinated bursting throughout reprogramming maintained differences in allelic accessibility (Fig 4B). Similarly, genes that maintained highly coordinated bursting throughout reprogramming always maintained similar allelic accessibility (Fig 4B). Furthermore, we show that allelic accessibility differences reduce upon becoming highly coordinated in iPSC after maintaining a semi-coordinated bursting throughout the other day points (Fig S3B). A similar trend was found for genes which switched from highly coordinated in MEF to semi-coordinated on other days (data not shown). Together, our analysis suggested a positive correlation between the coordination of allelic bursting and the similarity of allelic chromatin accessibility.

Next, we explored if allelic accessibility differences in semi-coordinated and independent genes are associated with the differential binding of transcription factors (TFs) between alleles of individual genes. To test this, we determined TFs binding scores of individual alleles of a gene using TOBIAS (Transcription factor Occupancy prediction By Investigation of ATAC-seq Signal). Interestingly, we found that many TFs had a significantly different binding score between individual alleles of semi-coordinated and independent genes, which was not the case for highly coordinated genes (Fig 5). Taken together, our analysis suggests that allelic accessibility differences in semi-coordinated or independent genes allow differential binding of certain TFs among alleles of genes, which in turn leads to semi-coordinated or independent transcriptional bursting.

Chromatin accessibility factors contribute to allelic bursting coordination

Next, we investigated if the degree of allelic coordination is dependent on chromatin accessibility factor enrichment. To explore this, we profiled the enrichment of different chromatin accessibility factors in the gene body and across 2 kb upstream of TSS and 2 kb downstream of TES of highly coordinated, semi-coordinated, and independent genes in MEF and iPSCs (Figs 6 and S4). Interestingly, we found that many accessibility factors such as H3.3, H3K36me3, H3K27ac, H3K4me3, H3K4me2, H3K4me1, and RNAPolIII-S2P are highly enriched on highly coordinated genes compared to the semi-

coordinated/independent genes in both MEF and iPSC (Figs 6 and S4). Moreover, in iPSC, we found that highly coordinated genes are enriched with chromatin remodeler BRD4 (Fig 6). H3K79me2, H3K9ac, and RNA PolIII-S5P were highly enriched to the highly coordinated genes in MEF cells (Figs 6 and S4). Next, we tested if switching of coordination pattern of genes in MEF to iPSC is associated with changes in the pattern of chromatin accessibility factor enrichment. We found that highly coordinated genes in MEF having higher enrichment of H3K36me3, H3K27ac, H3.3, and RNA PolIII-S2P compared to the semi-coordinated genes, which flip their pattern of enrichment in iPSC upon conversion of their coordination pattern (Fig S5A and B). Whereas genes that do not switch their coordination pattern in MEF to iPSC maintain the enrichment pattern of these marks (Fig S5A and B). Altogether, our analysis suggested that higher enrichment of these chromatin accessibility factors ensures highly coordinated allelic bursting. However, other chromatin modifications, such as H3K9me3 and CTCF, did not show such differences in enrichment between different categories of genes (Fig S4).

Next, we investigated if the chromatin states of highly coordinated, semi-coordinated, and independent genes are different. To explore this, we identified 10 chromatin states in both MEFs and iPSCs using the ChromHMM based on the seven histone marks (H3K27ac, H3K27me3, H3K36me3, H3K4me1, H3K4me2, H3K4me3, and H3K9me3) using the available ChIP-seq datasets (Fig S6). We annotated each state putatively based on related functional element enrichment. Next, we compared the enrichment of these different chromatin states in the gene body in the TSS/TES neighbourhood of highly coordinated, semi-coordinated, and independent genes. Interestingly, we found that chromatin states enriched around strong enhancers showed higher enrichment in highly coordinated genes than the semi-coordinated or independent genes in both MEF and iPSC (Fig S6A and B). Notably, independent genes were highly enriched with bivalent chromatin state compared to the highly coordinated genes in both MEF and iPSC (Fig S6A and B). On the other hand, enrichment of other chromatin states among highly coordinated, semi-coordinated, and independent genes did not exhibit much difference. Taken together, our analysis suggests that enrichment of enhancer elements in highly coordinated genes might play a role in coordinating the allelic transcriptional bursting.

Discussion

Since the discovery of reprogramming of somatic cells to pluripotent state in 2006, the underlying precise mechanistic aspect of reprogramming remains unknown (Takahashi & Yamanaka, 2006). To fill this gap, quantitatively understanding the transcriptional regulation of reprogramming-related genes is imperative. Emerging trends suggest that transcriptional regulation of two alleles of a gene is not always shared and can be independent, which in turn can lead to temporal expression heterogeneity (Finn et al, 2019;

kilobase million) of highly coordinated, semi-coordinated, and independent genes across different stages of reprogramming (Mann-Whitney *U* test: *P*-value < 0.0001; **** *P*-value < 0.01; **).

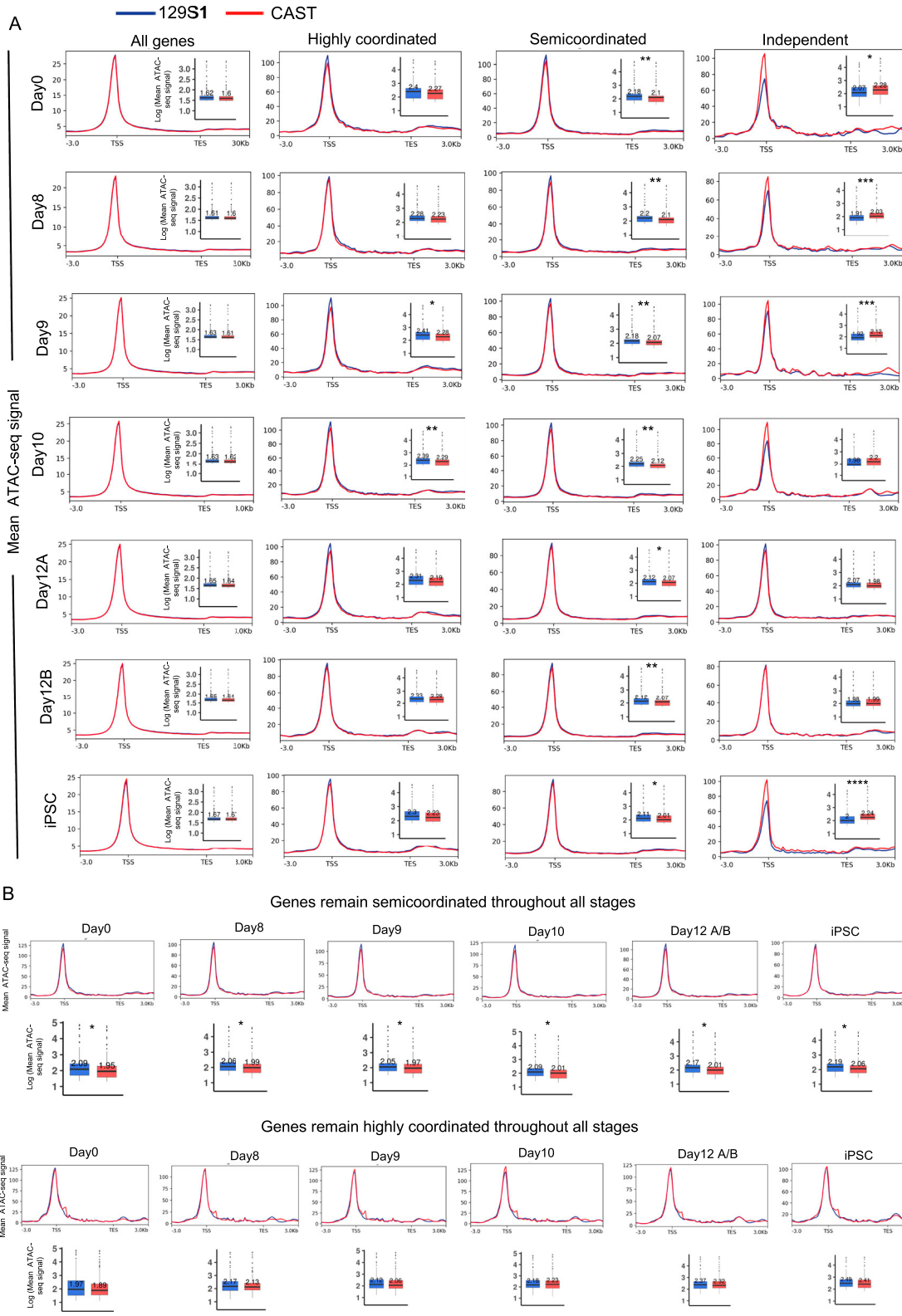


Figure 4. Comparison of allelic chromatin accessibility of highly coordinated, semi-coordinated, and independent genes.
(A) Quantitative analysis of allelic accessibility enrichment in the gene body and across 3 kb upstream of TSS and 3 kb downstream of TES of all autosomal genes, highly coordinated, semi-coordinated, and independent genes throughout different stages of reprogramming: day 0, day 8, day 10, day 12, and induced pluripotent stem cells. In the boxplots, the line inside each of the boxes denotes the median value, and the edges of each box represent 25% and 75% of dataset, respectively (Mann–Whitney U test:

Naik et al, 2021). Therefore, it is important to explore the allelic contribution of genes and their cooperativity to understand how cells fine-tune the optimal expression to bring developmental precision. To understand this aspect quantitatively, we have profiled genome-wide transcriptional burst kinetics at the allelic level and their relevance to cell state transition during iPSC reprogramming. We found that most of the autosomal genes exhibit bursty expression and have dynamic aRME across different stages of iPSC reprogramming (Fig 1), which is consistent with our previous reports in pre-gastrulation embryos (Naik et al, 2021). Importantly, we found that many genes involved in iPSC reprogramming pathways exhibit bursty expression (Fig 1). Interestingly, we found that burst frequency and burst size are highly similar between two alleles for most of the genes across reprogramming (Fig 2A and B). However, in terms of the cooperativity of allelic bursting, we found different patterns of allelic bursting: whereas most of the genes exhibited semi-coordinated allelic bursting, many genes showed highly coordinated allelic bursting (Fig 2C). On the other hand, few genes showed the independent nature of allelic bursting. It is worth mentioning that our analysis of allelic burst kinetics might be erroneous to some extent given the cellular heterogeneity during the iPSC reprogramming. Nevertheless, to reduce this kind of variability, we considered homogeneous populations of cells as clustered through the UMAP analysis. Separately, another limitation of our study is that we have not validated the burst kinetics deduced from the scRNA-seq data through other independent experiments. However, we must mention that many studies have leveraged single-cell RNA-seq technology coupled with allele-specific analysis to profile genome-wide transcriptional burst kinetics (Larsson et al, 2019a, 2019b; Ochiai et al, 2020; Johnsson et al, 2022). Notably, Ochiai et al demonstrated that burst kinetics measured through scRNA-seq data can be correlated with kinetics measured through single-molecule FISH (Ochiai et al, 2020). Furthermore, we demonstrate that highly coordinated genes have higher expression level compared to the semi-coordinated and independent genes, suggesting expression level is linked to the degree of coordination of allelic bursting (Fig 3D). In the future, a more precise investigation is necessary to disentangle the interconnection between the coordination of allelic bursting versus expression.

Next, we found that the degree of coordination of allelic transcriptional bursting is relevant to reprogramming pathways. We show that allelic bursting of many genes crucial to iPSC reprogramming occurs in a highly coordinated fashion (Fig 3). We found that genes related to translation, protein stability, protein folding, and RNA processing undergo highly coordinated allelic bursting in iPSC. Specially, ribosome biogenesis-related genes (e.g., *rpl7*, *nop58*, *nmd3*, *nifk*, etc.) become highly coordinated upon initiation of reprogramming (day 8 onwards) and remain highly coordinated through most of the intermediate stages and iPSC (Fig 3). Indeed, reprogramming of ribosome biogenesis is crucial for iPSC reprogramming and stem cell maintenance to enhance translational efficiency (Gabut et al, 2020; Hu, 2020). Moreover, pluripotent embryonic

stem cells bear a high density of inactive ribosomes to facilitate increased translation efficiency during their subsequent differentiation into different lineages (Sampath et al, 2008; Novak et al, 2012; You et al, 2015; Gabut et al, 2020). In addition, important genes like *nanog*, *dppa2*, *med28*, etc., which play important roles in iPSC reprogramming, showed highly coordinated allelic expression (Li et al, 2015). Importantly, many genes related to embryonic development and stem cell maintenance showed highly coordinated allelic bursting (Fig 3). Surprisingly, we observed that genes associated with mitochondrial cristae formation (e.g., *ndufa10*) exhibited highly coordinated allelic bursting in iPSC (Fig 3). It is known that iPSC reprogramming is associated with the metabolic shift from oxidative phosphorylation (OXPHOS) to glycolysis, which is accompanied by the formation of immature spherical mitochondria with less dense cristae in iPSCs in comparison to the highly elongated mitochondria with dense cristae network in MEF (Xu et al, 2013; Teslaa & Teitell, 2015; Seo et al, 2018). Although reprogramming to iPSC is associated with the shift from OXPHOS to glycolysis, iPSC remains dependent on mitochondrial metabolism for intermediate metabolites, which play an important role in epigenomic regulation to drive iPSC reprogramming or stem cell pluripotency (Todd et al, 2010; Carey et al, 2014; Moussaieff et al, 2015; Zhang et al, 2016, 2018; Zhou et al, 2016; Guitart et al, 2017; Fang et al, 2019). Therefore, it may be possible that highly coordinated bursting of these mitochondria-related genes is crucial for iPSC reprogramming. Notably, genes involved in nuclear pore formation shifted to highly coordinated state in iPSC. It has been demonstrated that the nuclear pore complex plays an important role in modulating pluripotency and reprogramming (Hansson et al, 2012; Yang et al, 2014). On the other hand, in a recent study, we demonstrated that developmental genes related to gastrulation undergo highly coordinated allelic bursting in pre-gastrulation mouse embryos (Naik et al, 2021). Taken together, we conclude that many genes crucial to reprogramming and development burst in a highly coordinated fashion. Possibly, the coordination of transcriptional bursting between the two alleles fine-tunes gene expression dosage to drive precise development or reprogramming.

Next, we show that epigenomic states contribute to the coordination between allelic transcriptional bursting. We demonstrate that the coordination of allelic bursting is linked to chromatin accessibility. We find that, whereas genes undergoing highly coordinated allelic bursting exhibit equivalent chromatin accessibility between alleles, alleles of semi-coordinated or independent genes have differential chromatin accessibility (Fig 4). Importantly, our analysis suggests that differential chromatin accessibility between alleles of semi-coordinated or independent genes leads to differential binding of certain transcription factors (Fig 5). Previous studies have shown that random chromatin accessibility contributes to transcriptional bursting by providing intermittent accessibility to the transcription factors (Brown et al, 2013; Nicolas et al, 2018; Chen et al, 2019; Fraser et al, 2021). Moreover, open chromatin states are believed to modulate transcriptional burst size and burst frequency (Bartman et al, 2016; Nicolas et al, 2018; Fraser et al, 2021; Bullock et al, 2022). Our analysis extends the support on the role of chromatin

P-value < 0.0001; **** *P*-value < 0.001; *** *P*-value < 0.01; ** and *P*-value < 0.05; *). (B) Allelic accessibility enrichment analysis in the gene body and across 3 kb upstream of TSS and 3 kb downstream of TES of genes that remain semicordinated or highly coordinated throughout the reprogramming (Mann–Whitney *U* test: *P*-value < 0.05; *).

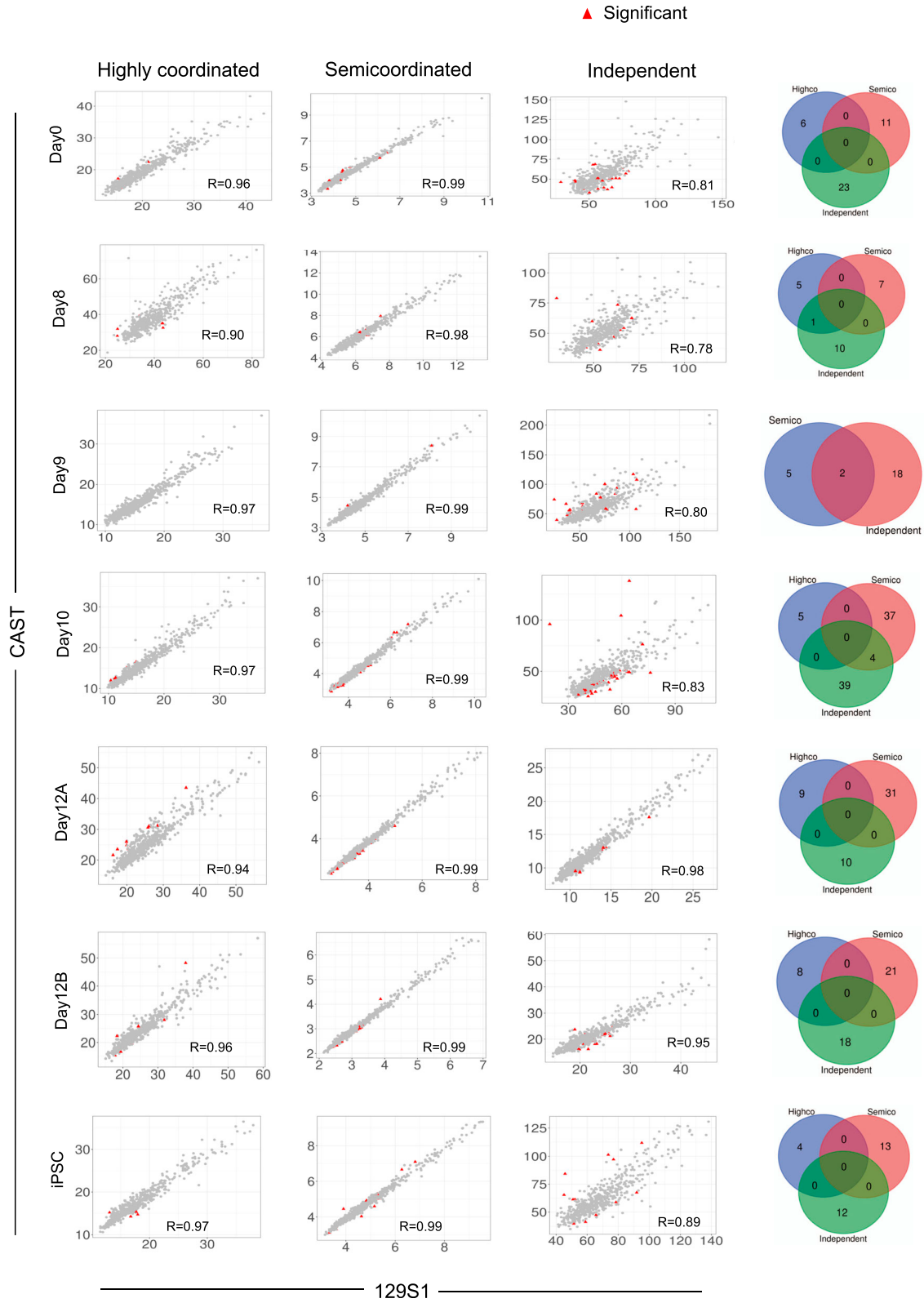


Figure 5. Binding kinetics of TFs between alleles correlate with the allelic bursting coordination. Left: Plots showing the correlation of TFs binding score between alleles of highly coordinated, semi-coordinated, and independent genes throughout different stages of reprogramming: day 0, day 8, day 10, day 12, and induced pluripotent stem cells. TFs that exhibit significant differences in binding between two alleles have been marked

accessibility on transcriptional bursting and, most importantly, sheds light on how allelic open chromatin dynamics is linked to allelic transcriptional burst kinetics. In the future, more extensive studies would help to gain deeper insights into the role of chromatin accessibility in mediating allelic bursting coordination.

On the other hand, we find that highly coordinated genes are enriched with important chromatin accessibility factors: H3K4me3, H3K36me3, H3K27ac, H3K4me1, BRD4, and histone variant H3.3 (Fig 6). In fact, H3K36me3, H3K27ac, along with BRD4, have been reported to play a role in orchestrating transcription and burst frequency (Nicolas et al, 2018; Ochiai et al, 2020; Sundarraj et al, 2021; Abe et al, 2022; Altendorfer et al, 2022; Pal et al, 2023). Interestingly, we report the implications of H3.3 in allelic transcriptional bursting for the first time. Moreover, we found that H3K79me2, H3K9ac, RNA PolII-S2P, and RNA PolII-S5P were also highly enriched in highly coordinated genes. Taken together, we propose that these chromatin-related factors play a crucial role in mediating the coordination of allelic bursting. In the future, analysis of allele-specific enrichment of these marks or factors in highly coordinated, semi-coordinated, and independent genes would provide better insight into the plausible regulatory link among allele-specific enrichment of active chromatin marks, burst frequency, and degree of coordination of bursting. Notably, many of these factors or marks have been reported to play an important role in iPSC reprogramming and pluripotency. BRD4 has been shown to play a crucial role in driving the late phase of iPSC reprogramming (Liu et al, 2014).

Occupancy of the H3.3 variant has been attributed to maintaining MEF-specific identity during the early stages of pluripotency reprogramming. However, towards the fag end of reprogramming, they have been conducive to determining and maintaining pluripotent cell fate (Fang et al, 2018). Moreover, H3.3 is found to be critical for early development as its depletion results in early embryonic lethality (Jang et al, 2015). Ablation of H3.3 has been implicated in reducing the overall chromatin accessibility in ESC (Tafessu et al, 2023). Overall, H3.3 enrichment is crucial in pluripotent cell maintenance and early development. Also, the enrichment of active chromatin marks, H3K27ac in enhancer, and H3K36me3 in gene bodies of mouse embryonic stem cells is important for their maintenance (Mikkelsen et al, 2007; Kim et al, 2018). Altogether, we conclude that these chromatin accessibility factors drive reprogramming through mediating allelic bursting coordination of genes involved in iPSC reprogramming to fine-tune the appropriate dosage of these genes (Fig 6). On the other hand, we show that chromatin states linked to enhancer element enrichment are highly enriched in highly coordinated genes (Fig S6). We conclude that enhancer element enrichment in highly coordinated genes might orchestrate the higher coordination of allelic bursting. Indeed, the previous report suggests that enhancer elements can contribute to coordinated transcriptional bursting (Fukaya et al, 2016). Broadly, our study provides fundamental insights into the implications of epigenomic states in fine-tuning of allelic dosage of genes to orchestrate cell fate specification and extends strong

support towards the role of epigenomic states in mediating transcriptional bursting (Fig 7).

Materials and Methods

Data acquisition

Single-cell RNA-seq, ATAC-seq datasets for MEF to iPSC reprogramming were retrieved from Gene Expression Omnibus under the following accessions: GSE153846 and GSE153844, respectively (Talon et al, 2021). The ChIP-seq datasets used for MEF and iPSC were retrieved from GSE87037 (Aldiri et al, 2017), GSE99592 (Fang et al, 2018), GSE33823 (Yildirim et al, 2011), and GSE90893 (Chronis et al, 2017).

Allele-specific expression and burst kinetics analysis

To obtain allelic read counts from scRNA-seq data, we performed allele-specific expression analysis following the pipeline as reported previously (Mandal et al, 2020; Naik et al, 2021, 2022). Briefly, we first constructed in silico CAST/EIJ and 129S1-specific parental genome by incorporating CAST/EIJ or 129S1-specific SNPs into the mm10 genome using variant calling file tool (Danecek et al, 2011). Variant calling file was downloaded from the mouse genome project (<https://www.sanger.ac.uk/science/data/mouse-genomes-project>). Next, we aligned RNA-seq reads into both parental genomes using STAR aligner (STAR-2.7.10a), allowing no multi-mapped reads (--outFilterMultimapNmax 1). We filtered out those genes for allele-specific read counts which had at least two informative SNPs and a minimum of three reads per SNP site. We took an average across SNPs to get gene-level allelic read counts. We normalized the allelic read counts by reads per kilobase million. Because low-expressed genes are dropout prone, we removed low-expressed genes from our analysis to avoid potential dropout effect (Kim et al, 2015; Santoni et al, 2017; Zhao et al, 2017). We considered only those genes, which were expressed in at least ~27% of cells and had mean expression >10 reads per kilobase million. Allelic ratio was calculated individually for each gene using the formula = (129S1/CAST reads) ÷ (129S1 + CAST reads). A gene was considered monoallelic if at least 95% of the allelic reads came from only one allele. We performed genome-wide allele-specific burst kinetics analysis using SCALE (Jiang et al, 2017). In brief, SCALE relies on Empirical Bayes Framework, which first classifies the genes into monoallelic, biallelic, and silent based on the allele-specific read counts and deduces the allelic burst kinetics based on the two-state model of transcription. It infers different burst kinetic parameters at the allelic level, such as burst frequency (K_{on}) and burst size (S/K_{off}), as described in the results. We excluded X-linked genes for our burst kinetics analysis. We performed SCALE analysis on each time point of the reprogramming separately.

by red triangles. Right: Venn diagram representing the comparison of TFs that exhibit significant differences in binding between two alleles across highly coordinated, semi-coordinated, and independent genes.

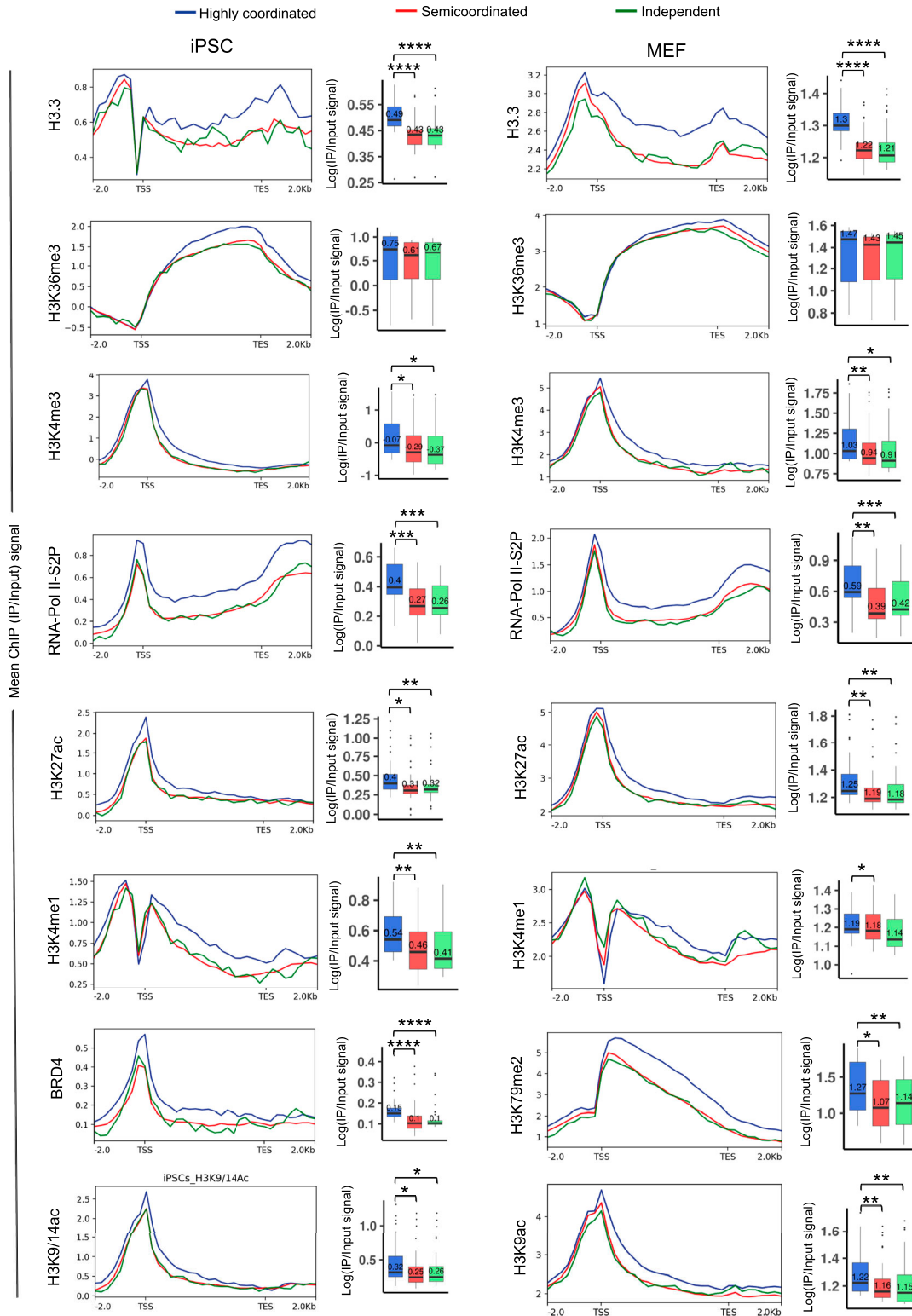


Figure 6. Correlation between occupancy of chromatin marks and degree of allelic bursting coordination.

Comparison of enrichment of different chromatin accessibility-related factors in the gene body and across 2 kb upstream of TSS and 2 kb downstream of TES of highly coordinated, semi-coordinated, and independent genes in MEF and induced pluripotent stem cells. In the boxplots, the line inside each of the boxes denotes the median value and the edges of each box represent 25% and 75% of dataset, respectively (Mann–Whitney *U* test: *P*-value < 0.0001, **** *P*-value < 0.001; *** *P*-value < 0.01; ** *P*-value < 0.05; *).



Reprogramming related genes show highly coordinated allelic bursting

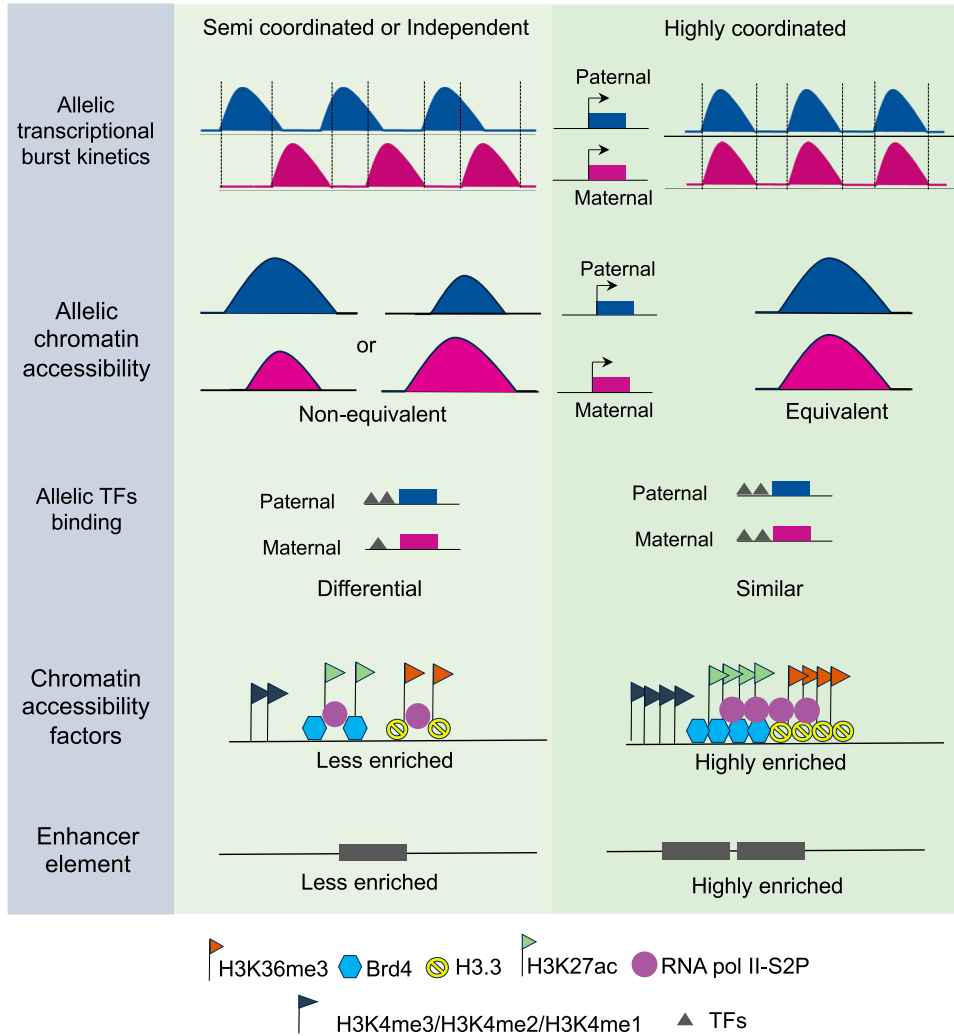


Figure 7. Epigenomic states mediate allelic bursting coordination.

Allelic chromatin accessibility is linked to allelic transcriptional bursting coordination. Genes with highly coordinated allelic bursting possess equivalent allelic chromatin accessibility. On the contrary, semi-coordinated or independent genes allelic chromatin accessibility differs. Importantly, highly coordinated genes are highly enriched with chromatin accessibility regulators and enhancer elements.

Gene ontology

Functional enrichment of different classes of genes was profiled using g:GOST from gProfiler (https://biit.cs.ut.ee/gprofiler_archive3/e108_eg55_p17/gost) with Benjamini–Hochberg FDR and selected the biological process having FDR < 0.05 from GO: BP (Raudvere et al, 2019).

Allelic ATAC-seq analysis

For allele-specific ATAC-seq analysis, first we created an “N-masked reference genome mm10” through substituting strain-specific (129S1/SvImj) and CAST/Eij) SNP position with “N” using

SNPsplit_genome_preparation (0.5.0) (Krueger & Andrews, 2016). Strain-specific SNPs were obtained as described above. Next, reads for all samples were mapped to this N-masked genome using Bowtie2 (Langmead & Salzberg, 2012). We removed duplicate reads and mitochondrial reads from our analysis. SNPsplit was then used to create allele-specific BAM files by segregating the aligned reads into two distinct alleles (129S1/SvImj and CAST/Eij). Bigwig files were generated from these allelic BAM files using deepTools (version 3.5.1) function bamCoverage (--binSize 100 --smoothLength 500 --normalizeUsing RPGC) (Ramirez et al, 2016). Allelic enrichment metaplots were generated for different gene categories using deepTools functions computeMatrix and plotProfile. Few genes were excluded from our analysis because of the lack of SNP for each

day: day 0–15 (of 1,745); day 8–13 (of 1,849); day 9–14 (of 1,906); day 10–19 (of 1,982); day 12A–21 (of 2,521); day 12B–23 (of 2,562); iPSC—14 (of 1,863).

ChIP-seq analysis

For analysis of chromatin mark enrichment in MEF and iPSC in different categories of genes, we analyzed available ChIP-seq data. For ChIP-seq analysis, the reads were mapped to either mm9 or mm10 genome using Bowtie tool (Langmead & Salzberg, 2012). We removed duplicate reads, mitochondrial reads from our analysis. Different enrichment metaplots were generated for different gene categories using deepTools (version 3.5.1) functions computeMatrix and plotProfile.

Transcription factor binding analysis using ATAC-seq footprinting

Transcription factor occupancy between individual alleles of genes was predicted using Transcription factor Occupancy prediction By Investigation of ATAC-seq Signal (TOBIAS, version 0.16.0) (Bentsen et al, 2020). Allelic ATAC-seq BAM files were first corrected for insertion bias of the Tn5 transposase using the command ATACCorrect for each gene category (highly coordinated, semi-coordinated, and independent) separately. Gene-wise allelic peak regions were identified through allelic peak calling using macs2 (v 2.2.7.1), and peaks were annotated using uropa (v 4.0.3). Next, footprinting scores for individual alleles were generated using the command ScoreBigwig using the ATACCorrect data.

Allelic footprint scores were matched to the curated list of JASPAR motifs using the command BINDetect for both alleles separately. Then the resulting scores of each TF motifs for both alleles were correlated using Pearson correlation method.

Identifications of chromatin states

Chromatin states in both MEFs and iPSCs were profiled using the ChromHMM (v 1.24) based on the seven chromatin marks (H3K27ac, H3K27me3, H3K36me3, H3K4me1, H3K4me2, H3K4me3, and H3K9me3) using the available ChIP-seq datasets (Ernst & Kellis, 2012). In brief, we first constructed binarized bam files of the chromatin marks required as input for ChromHMM using the binarizeBam function of ChromHMM v 1.24. Chromatin state models were generated with the LearnModel function with 10 states. Then we manually curated to assign putative functions to each state according to the combination of epigenomic marks. Finally, enrichment for each category of genes was shown across the states using the OverlapEnrichment function of ChromHMM.

Quantification and statistical analysis

All statistical analysis was performed using the R software (<https://www.R-project.org/>). Mann–Whitney two-sided *U* test was used for statistical significance analysis, and *P*-values < 0.05 was considered significant. For correlation analysis, the Pearson test was used.

Supplementary Information

Supplementary information is available at <https://doi.org/10.26508/lsa.202302337>.

Acknowledgements

This study is supported by DBT grant (BT/PR30399/BRB/10/1746/2018), DST-SERB (CRG/2019/003067), DBT-Ramalingaswamy fellowship (BT/RLF/Re-entry/05/2016), and Infosys Young Investigator Grant Award to S Gayen. AJ Naskar acknowledges Indian Institute of Science (IISc), Bangalore, for the fellowship. LS Bammidi acknowledges DST Women Scientist Award.

Author Contributions

P Ayyamperumal: conceptualization, data curation, formal analysis, validation, investigation, visualization, methodology, and writing—review and editing.

HC Naik: conceptualization, supervision, validation, visualization, methodology, and writing—review and editing.

AJ Naskar: conceptualization and writing—original draft.

LS Bammidi: conceptualization and writing—review and editing.

S Gayen: conceptualization, supervision, funding acquisition, investigation, and writing—original draft, review, and editing.

Conflict of Interest Statement

S Gayen is an advisor to Bangalore Bio Cluster.

References

- Abe K, Schauer T, Torres-Padilla ME (2022) Distinct patterns of RNA polymerase II and transcriptional elongation characterize mammalian genome activation. *Cell Rep* 41: 111865. doi:10.1016/j.celrep.2022.111865
- Aldiri I, Xu B, Wang L, Chen X, Hiler D, Griffiths L, Valentine M, Shirinifard A, Thiagarajan S, Sablauer A, et al (2017) The dynamic epigenetic landscape of the retina during development, reprogramming, and tumorigenesis. *Neuron* 94: 550–568.e10. doi:10.1016/j.neuron.2017.04.022
- Altendorfer E, Mochalova Y, Mayer A (2022) BRD4: A general regulator of transcription elongation. *Transcription* 13: 70–81. doi:10.1080/21541264.2022.2108302
- Bartman CR, Hsu SC, Hsiung CCS, Raj A, Blobel GA (2016) Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol Cell* 62: 237–247. doi:10.1016/j.molcel.2016.03.007
- Bauer M, Vidal E, Zorita E, Üresin N, Pinter SF, Filion GJ, Payer B (2021) Chromosome compartments on the inactive X guide TAD formation independently of transcription during X-reactivation. *Nat Commun* 12: 3499. doi:10.1038/S41467-021-23610-1
- Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, Fust A, Preussner J, Kuenne C, Braun T, et al (2020) ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat Commun* 11: 4267. doi:10.1038/s41467-020-18035-1
- Borel C, Ferreira PG, Santoni F, Delaneau O, Fort A, Popadin KY, Garieri M, Falconnet E, Ribaux P, Guipponi M, et al (2015) Biased allelic

- expression in human primary fibroblast single cells. *Am J Hum Genet* 96: 70–80. doi:[10.1016/j.ajhg.2014.12.001](https://doi.org/10.1016/j.ajhg.2014.12.001)
- Brown CR, Mao C, Falkovskaia E, Jurica MS, Boeger H (2013) Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biol* 11: e1001621. doi:[10.1371/JOURNAL.PBIO.1001621](https://doi.org/10.1371/JOURNAL.PBIO.1001621)
- Bullock ME, Moreno-Martinez N, Miller-Jensen K (2022) A transcriptional cycling model recapitulates chromatin-dependent features of noisy inducible transcription. *PLoS Comput Biol* 18: e1010152. doi:[10.1371/journal.pcbi.1010152](https://doi.org/10.1371/journal.pcbi.1010152)
- Carey BW, Finley LWS, Cross JR, Allis CD, Thompson CB (2014) Intracellular α -ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* 518: 413–416. doi:[10.1038/nature13981](https://doi.org/10.1038/nature13981)
- Cheloufi S, Elling U, Hopfgartner B, Jung YL, Murn J, Ninova M, Hubmann M, Badeaux AI, Euong Ang C, Tenen D, et al (2015) The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* 528: 218–224. doi:[10.1038/nature15749](https://doi.org/10.1038/nature15749)
- Chen LF, Lin YT, Gallegos DA, Hazlett MF, Gómez-Schiavon M, Yang MG, Kalmeta B, Zhou AS, Holtzman L, Gersbach CA, et al (2019) Enhancer histone acetylation modulates transcriptional bursting dynamics of neuronal activity-inducible genes. *Cell Rep* 26: 1174–1188.e5. doi:[10.1016/j.celrep.2019.01.032](https://doi.org/10.1016/j.celrep.2019.01.032)
- Chronis C, Fiziep P, Papp B, Butz S, Bonora G, Sabri S, Ernst J, Plath K (2017) Cooperative binding of transcription factors orchestrates reprogramming. *Cell* 168: 442–459.e20. doi:[10.1016/j.cell.2016.12.016](https://doi.org/10.1016/j.cell.2016.12.016)
- Chubb JR, Trcek T, Shenoy SM, Singer RH (2006) Transcriptional pulsing of a developmental gene. *Curr Biol* 16: 1018–1025. doi:[10.1016/j.cub.2006.03.092](https://doi.org/10.1016/j.cub.2006.03.092)
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al (2011) The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158. doi:[10.1093/bioinformatics/btr330](https://doi.org/10.1093/bioinformatics/btr330)
- Ernst J, Kellis M (2012) ChromHMM: Automating chromatin-state discovery and characterization. *Nat Methods* 9: 215–216. doi:[10.1038/nmeth.1906](https://doi.org/10.1038/nmeth.1906)
- Fang HT, El Farran CA, Xing QR, Zhang LF, Li H, Lim B, Loh YH (2018) Global H3.3 dynamic deposition defines its bimodal role in cell fate transition. *Nat Commun* 9: 1537. doi:[10.1038/s41467-018-03904-7](https://doi.org/10.1038/s41467-018-03904-7)
- Fang Y, Tang S, Li X (2019) Sirtuins in metabolic and epigenetic regulation of stem cells. *Trends Endocrinol Metab* 30: 177–188. doi:[10.1016/j.tem.2018.12.002](https://doi.org/10.1016/j.tem.2018.12.002)
- Finn EH, Pegoraro G, Brandão HB, Valton AL, Oomen ME, Dekker J, Mirny L, Misteli T (2019) Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* 176: 1502–1515.e10. doi:[10.1016/j.cell.2019.01.020](https://doi.org/10.1016/j.cell.2019.01.020)
- Fraser LCR, Dikdan RJ, Dey S, Singh A, Tyagi S (2021) Reduction in gene expression noise by targeted increase in accessibility at gene loci. *Proc Natl Acad Sci U S A* 118: e2018640118. doi:[10.1073/pnas.2018640118](https://doi.org/10.1073/pnas.2018640118)
- Fukaya T, Lim B, Levine M (2016) Enhancer control of transcriptional bursting. *Cell* 166: 358–368. doi:[10.1016/j.cell.2016.05.025](https://doi.org/10.1016/j.cell.2016.05.025)
- Gabut M, Bourdelais F, Durand S (2020) Ribosome and translational control in stem cells. *Cells* 9: 497. doi:[10.3390/cells9020497](https://doi.org/10.3390/cells9020497)
- Gayen S, Maclary E, Buttigieg E, Hinten M, Kalantry S (2015) A primary role for the tsix lncRNA in maintaining random X-chromosome inactivation. *Cell Rep* 11: 1251–1265. doi:[10.1016/j.celrep.2015.04.039](https://doi.org/10.1016/j.celrep.2015.04.039)
- Gayen S, Maclary E, Hinten M, Kalantry S (2016) Sex-specific silencing of X-linked genes by Xist RNA. *Proc Natl Acad Sci U S A* 113: E309–E318. doi:[10.1073/pnas.1515971113](https://doi.org/10.1073/pnas.1515971113)
- Gendrel AV, Marion-Poll L, Katoh K, Heard E (2016) Random monoallelic expression of genes on autosomes: Parallels with X-chromosome inactivation. *Semin Cell Dev Biol* 56: 100–110. doi:[10.1016/j.semcdb.2016.04.007](https://doi.org/10.1016/j.semcdb.2016.04.007)
- Generoso SF, Neguembor MV, Hershberg EA, Sadreyev RI, Kurimoto K, Yabuta Y, Ricci R, Audergon P, Bauer M, Saitou M, et al (2023) Cohesin controls X chromosome structure remodeling and X-reactivation during mouse iPSC-reprogramming. *Proc Natl Acad Sci U S A* 120: e2213810120. doi:[10.1073/PNAS.2213810120](https://doi.org/10.1073/PNAS.2213810120)
- Gregg C (2017) The emerging landscape of in vitro and in vivo epigenetic allelic effects. *F1000Research* 6: 2108. doi:[10.12688/f1000research.11491.1](https://doi.org/10.12688/f1000research.11491.1)
- Guitart AV, Panagopoulou TI, Villacreces A, Vukovic M, Sepulveda C, Allen L, Carter RN, van de Lagemaat LN, Morgan M, Giles P, et al (2017) Fumarate hydratase is a critical metabolic regulator of hematopoietic stem cell functions. *J Exp Med* 214: 719–735. doi:[10.1084/JEM.20161087](https://doi.org/10.1084/JEM.20161087)
- Hansson J, Rafiee MR, Reiland S, Polo JM, Gehring J, Okawa S, Huber W, Hochedlinger K, Krijgsvelde J (2012) Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep* 2: 1579–1592. doi:[10.1016/J.CELREP.2012.10.014](https://doi.org/10.1016/J.CELREP.2012.10.014)
- Hawkins KE, Joy S, Delhove JM, Kotiadis VN, Fernandez E, Fitzpatrick LM, Whiteford JR, King PJ, Bolanos JP, Duchon MR, et al (2016) NRF2 orchestrates the metabolic shift during induced pluripotent stem cell reprogramming. *Cell Rep* 14: 1883–1891. doi:[10.1016/j.celrep.2016.02.003](https://doi.org/10.1016/j.celrep.2016.02.003)
- Hu K (2020) Quick, coordinated and authentic reprogramming of ribosome biogenesis during iPSC reprogramming. *Cells* 9: 2484. doi:[10.3390/CELLS9112484](https://doi.org/10.3390/CELLS9112484)
- Jang CW, Shibata Y, Starmer J, Yee D, Magnuson T (2015) Histone H3.3 maintains genome integrity during mammalian development. *Genes Dev* 29: 1377–1392. doi:[10.1101/gad.264150.115](https://doi.org/10.1101/gad.264150.115)
- Jiang Y, Zhang NR, Li M (2017) SCALE: Modeling allele-specific gene expression by single-cell RNA sequencing. *Genome Biol* 18: 74. doi:[10.1186/s13059-017-1200-8](https://doi.org/10.1186/s13059-017-1200-8)
- Johnsson P, Ziegenhain C, Hartmanis L, Hendriks GJ, Hagemann-Jensen M, Reinius B, Sandberg R (2022) Transcriptional kinetics and molecular functions of long noncoding RNAs. *Nat Genet* 54: 306–317. doi:[10.1038/S41588-022-01014-1](https://doi.org/10.1038/S41588-022-01014-1)
- Kaltschmidt C, Greiner JFW, Kaltschmidt B (2021) The transcription factor nf- κ b in stem cells and development. *Cells* 10: 2042. doi:[10.3390/cells10082042](https://doi.org/10.3390/cells10082042)
- Kaur H, Rv P, Gayen S (2020) Dampened X-chromosomes in human pluripotent stem cells: Dampening or erasure of X-upregulation? *Chromosoma* 129: 111–113. doi:[10.1007/s00412-019-00717-5](https://doi.org/10.1007/s00412-019-00717-5)
- Kim JK, Kolodziejczyk AA, Ilicic T, Teichmann SA, Marioni JC (2015) Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun* 6: 8687. doi:[10.1038/ncomms9687](https://doi.org/10.1038/ncomms9687)
- Kim HS, Tan Y, Ma W, Merkurjev D, Destici E, Ma Q, Suter T, Ohgi K, Friedman M, Skowronska-Krawczyk D, et al (2018) Pluripotency factors functionally premark cell-type-restricted enhancers in ES cells. *Nature* 556: 510–514. doi:[10.1038/s41586-018-0048-8](https://doi.org/10.1038/s41586-018-0048-8)
- Krueger F, Andrews SR (2016) SNPsplite: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Research* 5: 1479. doi:[10.12688/F1000RESEARCH.9037.2](https://doi.org/10.12688/F1000RESEARCH.9037.2)
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359. doi:[10.1038/NMETH.1923](https://doi.org/10.1038/NMETH.1923)
- Larson DR (2011) What do expression dynamics tell us about the mechanism of transcription? *Curr Opin Genet Dev* 21: 591–599. doi:[10.1016/j.gde.2011.07.010](https://doi.org/10.1016/j.gde.2011.07.010)
- Larsson AJM, Coucoravas C, Sandberg R, Reinius B (2019a) X-chromosome upregulation is driven by increased burst frequency. *Nat Struct Mol Biol* 26: 963–969. doi:[10.1038/s41594-019-0306-y](https://doi.org/10.1038/s41594-019-0306-y)
- Larsson AJM, Johnsson P, Hagemann-Jensen M, Hartmanis L, Faridani OR, Reinius B, Segerstolpe Å, Rivera CM, Ren B, Sandberg R (2019b)

- Genomic encoding of transcriptional burst kinetics. *Nature* 565: 251–254. doi:[10.1038/s41586-018-0836-1](https://doi.org/10.1038/s41586-018-0836-1)
- Li L, Walsh RM, Wagh V, James MF, Beauchamp RL, Chang YS, Gusella JF, Hochedlinger K, Ramesh V (2015) Mediator subunit MED28 is essential for mouse peri-implantation development and pluripotency. *PLoS One* 10: e0140192. doi:[10.1371/journal.pone.0140192](https://doi.org/10.1371/journal.pone.0140192)
- Little SC, Tikhonov M, Gregor T (2013) Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* 154: 789–800. doi:[10.1016/j.cell.2013.07.025](https://doi.org/10.1016/j.cell.2013.07.025)
- Liu L, Xu Y, He M, Zhang M, Cui F, Lu L, Yao M, Tian W, Benda C, Zhuang Q, et al (2014) Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. *Cell Stem Cell* 15: 574–588. doi:[10.1016/j.stem.2014.09.018](https://doi.org/10.1016/j.stem.2014.09.018)
- Mandal S, Chandel D, Kaur H, Majumdar S, Arava M, Gayen S (2020) Single-cell analysis reveals partial reactivation of X chromosome instead of chromosome-wide dampening in naive human pluripotent stem cells. *Stem Cell Rep* 14: 745–754. doi:[10.1016/j.stemcr.2020.03.027](https://doi.org/10.1016/j.stemcr.2020.03.027)
- Mathieu J, Ruohola-Baker H (2017) Metabolic remodeling during the loss and acquisition of pluripotency. *Development* 144: 541–551. doi:[10.1242/dev.128389](https://doi.org/10.1242/dev.128389)
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448: 553–560. doi:[10.1038/nature06008](https://doi.org/10.1038/nature06008)
- Moussaieff A, Rouleau M, Kitsberg D, Cohen M, Levy G, Barasch D, Nemirovski A, Shen-Orr S, Laevsky I, Amit M, et al (2015) Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab* 21: 392–402. doi:[10.1016/j.cmet.2015.02.002](https://doi.org/10.1016/j.cmet.2015.02.002)
- Naik HC, Hari K, Chandel D, Mandal S, Jolly MK, Gayen S (2021) Semicordinated allelic-bursting shape dynamic random monoallelic expression in pregastrulation embryos. *iScience* 24: 102954. doi:[10.1016/j.isci.2021.102954](https://doi.org/10.1016/j.isci.2021.102954)
- Naik HC, Hari K, Chandel D, Jolly MK, Gayen S (2022) Single-cell analysis reveals X upregulation is not global in pre-gastrulation embryos. *iScience* 25: 104465. doi:[10.1016/j.isci.2022.104465](https://doi.org/10.1016/j.isci.2022.104465)
- Naik H, Chandel D, Majumdar S, Arava M, Baro R, Bv H, Hari K, Parichitran A, Manhas A, Jolly M, et al (2023) Lineage-specific dynamics of erasure of X-upregulation during inactive-X reactivation. *BioRxiv*. doi:[10.1101/2020.12.23.424181](https://doi.org/10.1101/2020.12.23.424181) (Preprint posted March 28, 2023).
- Nicolas D, Zoller B, Suter DM, Naef F (2018) Modulation of transcriptional burst frequency by histone acetylation. *Proc Natl Acad Sci U S A* 115: 7153–7158. doi:[10.1073/pnas.1722330115](https://doi.org/10.1073/pnas.1722330115)
- Novak A, Amit M, Ziv T, Segev H, Fishman B, Admon A, Itskovitz-Eldor J (2012) Proteomics profiling of human embryonic stem cells in the early differentiation stage. *Stem Cell Rev Rep* 8: 137–149. doi:[10.1007/s12015-011-9286-y](https://doi.org/10.1007/s12015-011-9286-y)
- Ochiai H, Hayashi T, Umeda M, Yoshimura M, Harada A, Shimizu Y, Nakano K, Saitoh N, Liu Z, Yamamoto T, et al (2020) Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Sci Adv* 6: eaaz6699. doi:[10.1126/sciadv.aaz6699](https://doi.org/10.1126/sciadv.aaz6699)
- Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, Cahan P, Marcarci BO, Unternaehrer J, Gupta PB, et al (2012) Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483: 598–602. doi:[10.1038/nature10953](https://doi.org/10.1038/nature10953)
- Onuchic V, Lurie E, Carrero I, Pawliczek P, Patel RY, Rozowsky J, Galeev T, Huang Z, Altshuler RC, Zhang Z, et al (2018) Allele-specific epigenome maps reveal sequence-dependent stochastic switching at regulatory loci. *Science* 361: eaar3146. doi:[10.1126/science.aar3146](https://doi.org/10.1126/science.aar3146)
- Padovan-Merhar O, Nair GP, Biaesch AG, Mayer A, Scarfone S, Foley SW, Wu AR, Churchman LS, Singh A, Raj A (2015) Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol Cell* 58: 339–352. doi:[10.1016/j.molcel.2015.03.005](https://doi.org/10.1016/j.molcel.2015.03.005)
- Pal D, Patel M, Boulet F, Sundarraj J, Grant OA, Branco MR, Basu S, Santos SDM, Zabet NR, Scaffidi P, et al (2023) H4K16ac activates the transcription of transposable elements and contributes to their cis-regulatory function. *Nat Struct Mol Biol* 30: 935–947. doi:[10.1038/s41594-023-01016-5](https://doi.org/10.1038/s41594-023-01016-5)
- Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, Borkent M, Apostolou E, Alaei S, Cloutier J, et al (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151: 1617–1632. doi:[10.1016/j.cell.2012.11.039](https://doi.org/10.1016/j.cell.2012.11.039)
- Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* 135: 216–226. doi:[10.1016/j.cell.2008.09.050](https://doi.org/10.1016/j.cell.2008.09.050)
- Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* 4: e309. doi:[10.1371/journal.pbio.0040309](https://doi.org/10.1371/journal.pbio.0040309)
- Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T (2016) deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44: W160–W165. doi:[10.1093/nar/gkw257](https://doi.org/10.1093/nar/gkw257)
- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019) G: Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* 47: W191–W198. doi:[10.1093/nar/gkz369](https://doi.org/10.1093/nar/gkz369)
- Reinius B, Sandberg R (2015) Random monoallelic expression of autosomal genes: Stochastic transcription and allele-level regulation. *Nat Rev Genet* 16: 653–664. doi:[10.1038/nrg3888](https://doi.org/10.1038/nrg3888)
- Reinius B, Mold JE, Ramsköld D, Deng Q, Johnsson P, Michaëlsson J, Frisén J, Sandberg R (2016) Analysis of allelic expression patterns in clonal somatic cells by single-cell RNA-seq. *Nat Genet* 48: 1430–1435. doi:[10.1038/ng.3678](https://doi.org/10.1038/ng.3678)
- RV P, Sundaresh A, Karunyaa M, Arun A, Gayen S (2021) Autosomal clonal monoallelic expression: Natural or artifactual? *Trends Genet* 37: 206–211. doi:[10.1016/j.tig.2020.10.011](https://doi.org/10.1016/j.tig.2020.10.011)
- Saiba R, Arava M, Gayen S (2018) Dosage compensation in human pre-implantation embryos: X-chromosome inactivation or dampening? *EMBO Rep* 19: e46294. doi:[10.15252/embr.201846294](https://doi.org/10.15252/embr.201846294)
- Samanta MK, Gayen S, Harris C, Maclary E, Murata-Nakamura Y, Malcore RM, Porter RS, Garay PM, Vallianatos CN, Samollow PB, et al (2022) Activation of Xist by an evolutionarily conserved function of KDM5C demethylase. *Nat Commun* 13: 2602. doi:[10.1038/S41467-022-30352-1](https://doi.org/10.1038/S41467-022-30352-1)
- Sampath P, Pritchard DK, Pabon L, Reinecke H, Schwartz SM, Morris DR, Murry CE (2008) A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2: 448–460. doi:[10.1016/j.stem.2008.03.013](https://doi.org/10.1016/j.stem.2008.03.013)
- Santoni FA, Stamoulis G, Garieri M, Falconnet E, Ribaux P, Borel C, Antonarakis SE (2017) Detection of imprinted genes by single-cell allele-specific gene expression. *Am J Hum Genet* 100: 444–453. doi:[10.1016/j.ajhg.2017.01.028](https://doi.org/10.1016/j.ajhg.2017.01.028)
- Sarkar MK, Gayen S, Kumar S, Maclary E, Buttigieg E, Hinten M, Kumari A, Harris C, Sado T, Kalantry S (2015) An Xist-activating antisense RNA required for X-chromosome inactivation. *Nat Commun* 6: 8564. doi:[10.1038/ncomms9564](https://doi.org/10.1038/ncomms9564)
- Schröter F, Adjaye J (2014) The proteasome complex and the maintenance of pluripotency: Sustain the fate by mopping up? *Stem Cell Res Ther* 5: 24–29. doi:[10.1186/scrt413](https://doi.org/10.1186/scrt413)
- Seo BJ, Yoon SH, Do JT (2018) Mitochondrial dynamics in stem cells and differentiation. *Int J Mol Sci* 19: 3893. doi:[10.3390/IJMS19123893](https://doi.org/10.3390/IJMS19123893)
- Sundarraj J, Taylor GCA, von Kriegsheim A, Pradeepa MM (2021) H3K36me3 and PSIP1/LEDGF associate with several DNA repair proteins,

- suggesting their role in efficient DNA repair at actively transcribing loci. *Wellcome Open Res* 2: 83. doi:[10.12688/wellcomeopenres.11589.4](https://doi.org/10.12688/wellcomeopenres.11589.4)
- Szutorisz H, Georgiou A, Tora L, Dillon N (2006) The proteasome restricts permissive transcription at tissue-specific gene loci in embryonic stem cells. *Cell* 127: 1375–1388. doi:[10.1016/j.cell.2006.10.045](https://doi.org/10.1016/j.cell.2006.10.045)
- Tafessu A, O'Hara R, Martire S, Dube AL, Saha P, Gant VU, Banaszynski LA (2023) H3.3 contributes to chromatin accessibility and transcription factor binding at promoter-proximal regulatory elements in embryonic stem cells. *Genome Biol* 24: 25. doi:[10.1186/s13059-023-02867-3](https://doi.org/10.1186/s13059-023-02867-3)
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663–676. doi:[10.1016/J.CELL.2006.07.024](https://doi.org/10.1016/J.CELL.2006.07.024)
- Takahashi K, Yamanaka S (2016) A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol* 17: 183–193. doi:[10.1038/nrm.2016.8](https://doi.org/10.1038/nrm.2016.8)
- Talon I, Janiszewski A, Theeuwes B, Lefevre T, Song J, Bervoets G, Vanheer L, De Geest N, Poovathingal S, Allsop R, et al (2021) Enhanced chromatin accessibility contributes to X chromosome dosage compensation in mammals. *Genome Biol* 22: 302. doi:[10.1186/s13059-021-02518-5](https://doi.org/10.1186/s13059-021-02518-5)
- Teslaa T, Teitell MA (2015) Pluripotent stem cell energy metabolism: An update. *EMBO J* 34: 138–153. doi:[10.15252/embj.201490446](https://doi.org/10.15252/embj.201490446)
- Todd LR, Damin MN, Gomathinayagam R, Horn SR, Means AR, Sankar U (2010) Growth factor ero1-like modulates Drp1 to preserve mitochondrial dynamics and function in mouse embryonic stem cells. *Mol Biol Cell* 21: 1225–1236. doi:[10.1091/MBE.09-11-0937/ASSET/IMAGES/LARGE/ZMK0071094110007](https://doi.org/10.1091/MBE.09-11-0937/ASSET/IMAGES/LARGE/ZMK0071094110007)
- Tohyama S, Tanosaki S, Someya S, Fujita J, Fukuda K (2017) Manipulation of pluripotent stem cell metabolism for clinical application. *Curr Stem Cell Rep* 3: 28–34. doi:[10.1007/s40778-017-0073-9](https://doi.org/10.1007/s40778-017-0073-9)
- Xu X, Duan S, Yi F, Ocampo A, Liu GH, Izpisua Belmonte JC (2013) Mitochondrial regulation in pluripotent stem cells. *Cell Metab* 18: 325–332. doi:[10.1016/j.cmet.2013.06.005](https://doi.org/10.1016/j.cmet.2013.06.005)
- Yan P, Ren J, Zhang W, Qu J, Liu GH (2020) Protein quality control of cell stemness. *Cell Regen* 9: 22. doi:[10.1186/s13619-020-00064-2](https://doi.org/10.1186/s13619-020-00064-2)
- Yang J, Cai N, Yi F, Liu GH, Qu J, Izpisua Belmonte JC (2014) Gating pluripotency via nuclear pores. *Trends Mol Med* 20: 1–7. doi:[10.1016/j.molmed.2013.10.003](https://doi.org/10.1016/j.molmed.2013.10.003)
- Yildirim E, Sadreyev RI, Pinter SF, Lee JT (2011) X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. *Nat Struct Mol Biol* 19: 56–61. doi:[10.1038/nsmb.2195](https://doi.org/10.1038/nsmb.2195)
- You KT, Park J, Kim VN (2015) Role of the small subunit processome in the maintenance of pluripotent stem cells. *Genes Dev* 29: 2004–2009. doi:[10.1101/GAD.267112.115](https://doi.org/10.1101/GAD.267112.115)
- Zhang H, Badur MG, Divakaruni AS, Parker SJ, Jäger C, Hiller K, Murphy AN, Metallo CM (2016) Distinct metabolic states can support self-renewal and lipogenesis in human pluripotent stem cells under different culture conditions. *Cell Rep* 16: 1536–1547. doi:[10.1016/J.CELREP.2016.06.102](https://doi.org/10.1016/J.CELREP.2016.06.102)
- Zhang J, Zhao J, Dahan P, Lu V, Zhang C, Li H, Teitell MA (2018) Metabolism in pluripotent stem cells and early mammalian development. *Cell Metab* 27: 332–338. doi:[10.1016/J.CMET.2018.01.008](https://doi.org/10.1016/J.CMET.2018.01.008)
- Zhao D, Lin M, Pedrosa E, Lachman HM, Zheng D (2017) Characteristics of allelic gene expression in human brain cells from single-cell RNA-seq data analysis. *BMC Genomics* 18: 860. doi:[10.1186/s12864-017-4261-x](https://doi.org/10.1186/s12864-017-4261-x)
- Zhou G, Meng S, Li Y, Ghebre YT, Cooke JP (2016) Optimal ROS signaling is critical for nuclear reprogramming. *Cell Rep* 15: 919–925. doi:[10.1016/J.CELREP.2016.03.084](https://doi.org/10.1016/J.CELREP.2016.03.084)



License: This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).