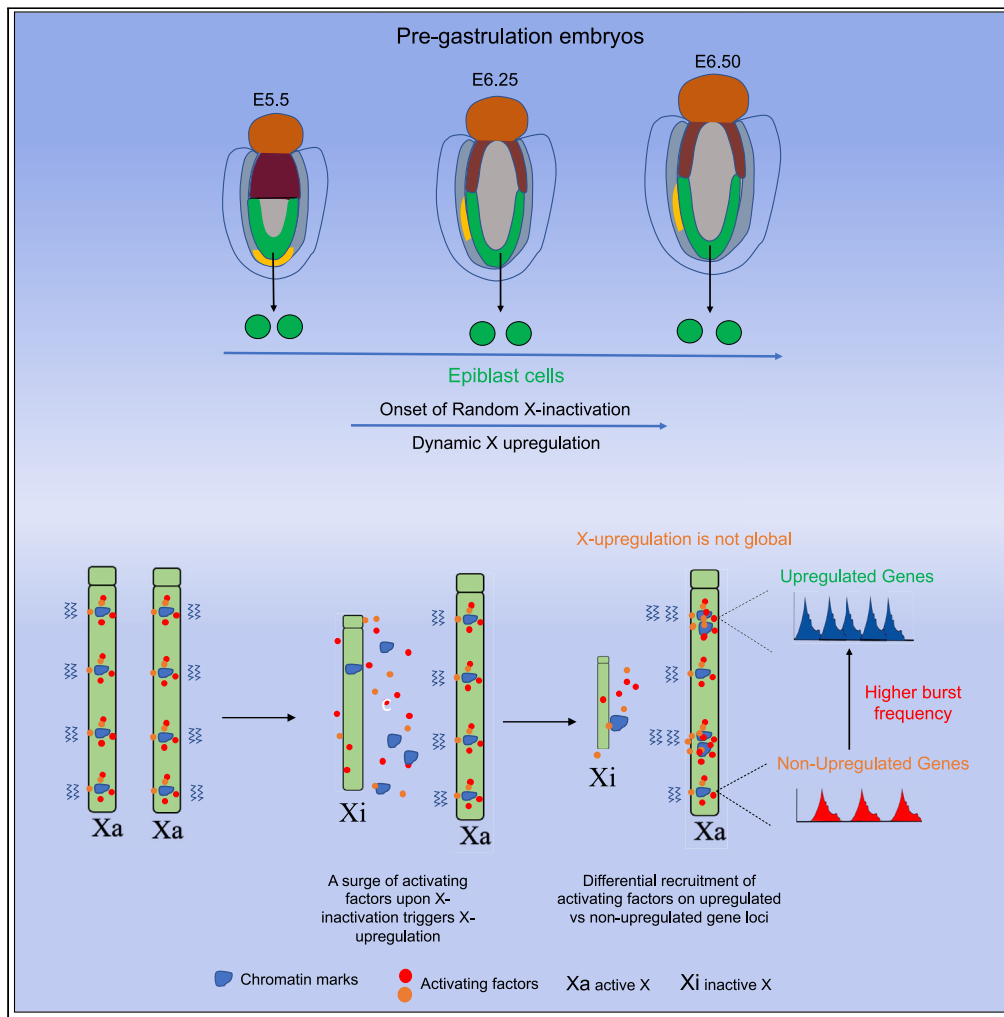# iScience

Article

Single-cell analysis reveals X upregulation is not global in pre-gastrulation embryos

Hemant Chandru
Naik, Kishore Hari,
Deepshikha
Chandel, Mohit
Kumar Jolly,
Srimonta Gayen

srimonta@iisc.ac.in

Highlights

X-upregulation coincides
with X chromosome
inactivation in pre-
gastrulation embryos

X-upregulation is not
chromosome-wide like
X-inactivation

Upregulated genes have
enhanced burst frequency
and are enriched with
activating marks

A surge of activating
factors on X-inactivation
triggers X-upregulation

iScience

Article

Single-cell analysis reveals X upregulation is not global in pre-gastrulation embryos

Hemant Chandru Naik,[1] Kishore Hari,[2] Deepshikha Chandel,[1] Mohit Kumar Jolly,[2] and Srimonta Gayen[1,3,*]

SUMMARY

In mammals, transcriptional inactivation of one X chromosome in female compensates for the dosage of X-linked gene expression between the sexes. Additionally, it is believed that the upregulation of active X chromosome in male and female balances the dosage of X-linked gene expression relative to autosomal genes, as proposed by Ohno. However, the existence of X chromosome upregulation (XCU) remains controversial. Here, we have profiled gene-wise dynamics of XCU in pre-gastrulation mouse embryos at single-cell level and found that XCU is dynamically linked with X chromosome inactivation (XCI); however, XCU is not global like XCI. Moreover, we show that upregulated genes are enriched with activating marks and have enhanced burst frequency. Finally, our In-silico model predicts that recruitment probabilities of activating factors and a surge of these factors upon X-inactivation trigger XCU. Altogether, our study provides significant insight into the gene-wise dynamics and mechanistic basis of XCU during early development and extends support for Ohno's hypothesis.

INTRODUCTION

In therian mammals, sex is determined by the sex chromosomes: XX (female) and XY (male). The X and Y evolved from a pair of autosomal homologs around 166–180 million years ago through the acquisition of male determining *Sry* gene on one of the chromosomes, which became *de facto* proto-Y (Cortez et al., 2014; Veyrunes et al., 2008). During evolution, the acquisition of male genes on the Y led to the loss of recombination, resulting progressive degradation of the Y chromosome. Degradation of Y created a dosage imbalance between X and autosomal genes in males and between the sexes (Graves, 2016). In 1967, Ohno hypothesized that the evolution of dosage compensation happened through two steps: X chromosome in male cells was upregulated to 2-fold to correct the dosage imbalance related to Y degradation (Ohno, 1967). Subsequently, this X chromosome upregulation (XCU) was inherited in females and thereby introduced an extra dosage of the X chromosome in female cells. Therefore, to restore optimal dosage from X chromosome in female cells, the evolution of X chromosome inactivation (XCI) happened, a process that silent one of the X chromosome in female mammals (Lyon, 1961). However, Ohno's hypothesis was not accepted well for a long time owing to the lack of proper experimental evidence. The first evidence of X-upregulation came through the studies based on microarray analysis; however, subsequently, it was challenged through RNA-seq based analysis (Nguyen and Disteche, 2006; Gupta et al., 2006; Johnston et al., 2008; Lin et al., 2007; Talebizadeh et al., 2006; Xiong et al., 2010). Down the line, several independent studies came out both supporting and refuting Ohno's hypothesis (Chen and Zhang, 2016; Moreira de Mello et al., 2017; Deng et al., 2011, 2013; Julien et al., 2012; Kharchenko et al., 2011; Larsson et al., 2019; Li et al., 2017b; Lin et al., 2011; Sangrithi et al., 2017). We have recently shown that there is, indeed, the presence of upregulated active-X chromosome (X_{2a}) in human pluripotent stem cells *in vitro* (Mandal et al., 2020). However, the dynamics of XCU at the onset of XCI during early embryonic development remains poorly understood. On the other hand, the extent of XCU, i.e., whether XCU is chromosome-wide like XCI or restricted to specific genes, remains unknown. Many studies implicated that XCU might be global, though direct evidence showing gene-wise dynamics of upregulation is not well understood (Deng et al., 2011, 2013). On contrary, several studies implicated that XCU affects dosage-sensitive genes such as components of macromolecular complexes, signal transduction pathways, or encoding for transcription factors (Pessia et al., 2012, 2014). To get better insight into this, here we have explored gene-wise dynamics of XCU in pre-gastrulation mouse embryos at single-cell level through allele-specific

single-cell RNA-seq analysis. We found that XCU is neither chromosome-wide nor restricted to dosage-sensitive genes only. Next, we have explored the mechanistic basis of why some genes are upregulated while others are not from the same active X chromosome and how XCU is linked to the XCI.

RESULTS

Dynamic active-X upregulation upon random XCI in epiblast cells of pre-gastrulation embryos

In female pre-gastrulation mouse embryos, while extraembryonic cells harbor imprinted inactive X, epiblast cells are at the onset of random XCI (Gayen et al., 2015, 2016; C. Harris et al., 2019; Maclary et al., 2017; Sarkar et al., 2015). We investigated the status of XCU in these different lineages of pre-gastrulation embryos by performing allelic/non-allelic gene expression analysis using the available scRNA-seq dataset of E5.5, E6.25, and E6.5 hybrid mouse embryos (Cheng et al., 2019) (Figure 1A). These embryos were derived from two divergent mouse strains (C57BL/6J and CAST/EiJ) and therefore harbored polymorphic sites across the genome, allowing us to profile gene expression with allelic resolution (Figure 1A). First, we classified the cells of E5.5, E6.25, and E6.5 mouse embryos into the three lineages: epiblast (EPI), extraembryonic ectoderm (ExE), and visceral endoderm (VE) based on our previous work (see detail in STAR Methods) (Naik et al., 2021) (Figure 1A). Next, we categorized the cells of different lineages of the female embryos based on their XCI status: cells with no XCI (X_aX_a), partial or undergoing XCI (X_aX_p), and complete XCI (X_aX_i) through profiling the fraction of maternal allele expression (Figure 1B). We found that in the EPI lineage lot of cells belong to X_aX_p/X_aX_a category, indicating that these cells are the onset of random X-inactivation. In contrast, cells of VE/ExE lineage mostly belong to X_aX_i category, indicating the establishment of imprinted X-inactivation (Figure 1B). As expected, autosomal genes showed almost equivalent paternal/maternal allele expression, thus validating our allele-specific analysis (Figure S1A). Next, to check the upregulation dynamics, we profiled X:A ratio in the individual cells of different stages/lineages of embryos. If a diploid female cell with an inactive-X (X_aX_i) poses upregulated active X, the expected X:A ratio should be more than 0.5 and closer to 1. We found that X:A ratio of X_aX_p/X_aX_i cells is always >0.5 and close to one despite XCI, indicating dynamic X-upregulation from the active-X chromosome X_aX_p/X_aX_i cells (Figure 1C). Similarly, male cells had X:A ratio >0.5 and close to 1, suggesting the presence of upregulated X chromosome (Figure 1C). Next, to validate the presence of XCU in more accurately, we compared the allelic expression pattern from autosome and X chromosome in individual cells of each lineage. Indeed, we found that the active X expression is always significantly higher than the allelic expression of autosomal genes in X_aX_p/X_aX_i cells, reflecting the upregulation of X-linked gene expression from the active X chromosome (Figure 1D and Tables S1 and S2). On the other hand, there were no significant differences in active Xs and autosomal allelic expression in X_aX_a cells of epiblast lineage, suggesting no upregulation of X chromosome in the absence of XCI (Figure 1D and Table S2). As expected, X chromosome in male cells also showed significantly higher expression than each allele of autosomes (Figures 1D, S1B, and S1C; Table S2). Altogether, these analyses suggested (a) dynamic active XCU on the initiation of random XCI in female epiblast cells, (b) presence of upregulated X-Chr. in female VE and ExE cells with imprinted X inactivation, and (c) presence of upregulated X-Chr. in different lineages of male cells of pre-gastrulation embryos as well.

X chromosome upregulation is not global

Our X:A analysis in female pre-gastrulation embryos revealed that the X:A ratio of XaXi cells is consistently lower compared to the XaXa cells (Figure 1C). This data hinted that the X-upregulation of X-linked genes in XaXi cells is partial or all genes do not undergo upregulation. To explore this further, we investigated whether X chromosome upregulation occurs globally or is restricted to certain genes. To test this, we profiled gene-wise dynamics of XCU by comparing the expression of X-linked genes from the active X chromosome of X_aX_i cells with the same active allele of X_aX_a cells in EPI E6.5 (Figure 2A). If active allele of a gene is upregulated in X_aX_i cells, it will show increased expression from the active allele in X_aX_i cells compared to the same active allele of X_aX_a cells. We found that while many X-linked genes showed increased expression from the active-X allele in X_aX_i cells compared to the X_aX_a, a significant number of genes did not show such increased expression, suggesting that all genes do not undergo upregulation (Figure 2A). Moreover, while some genes showed robust upregulation from the active-X allele, the others were moderately upregulated. Altogether, this result suggested that X-upregulation is not global or does not occur chromosome-wide. On the other hand, surprisingly we found that an adequate number of upregulated X-linked genes showed allele-specificity as they showed upregulation from either C57 or CAST as an active-X, suggesting that the regulation of upregulation of active allele may occur in the parent of origin-specific manner (Figure 2B).
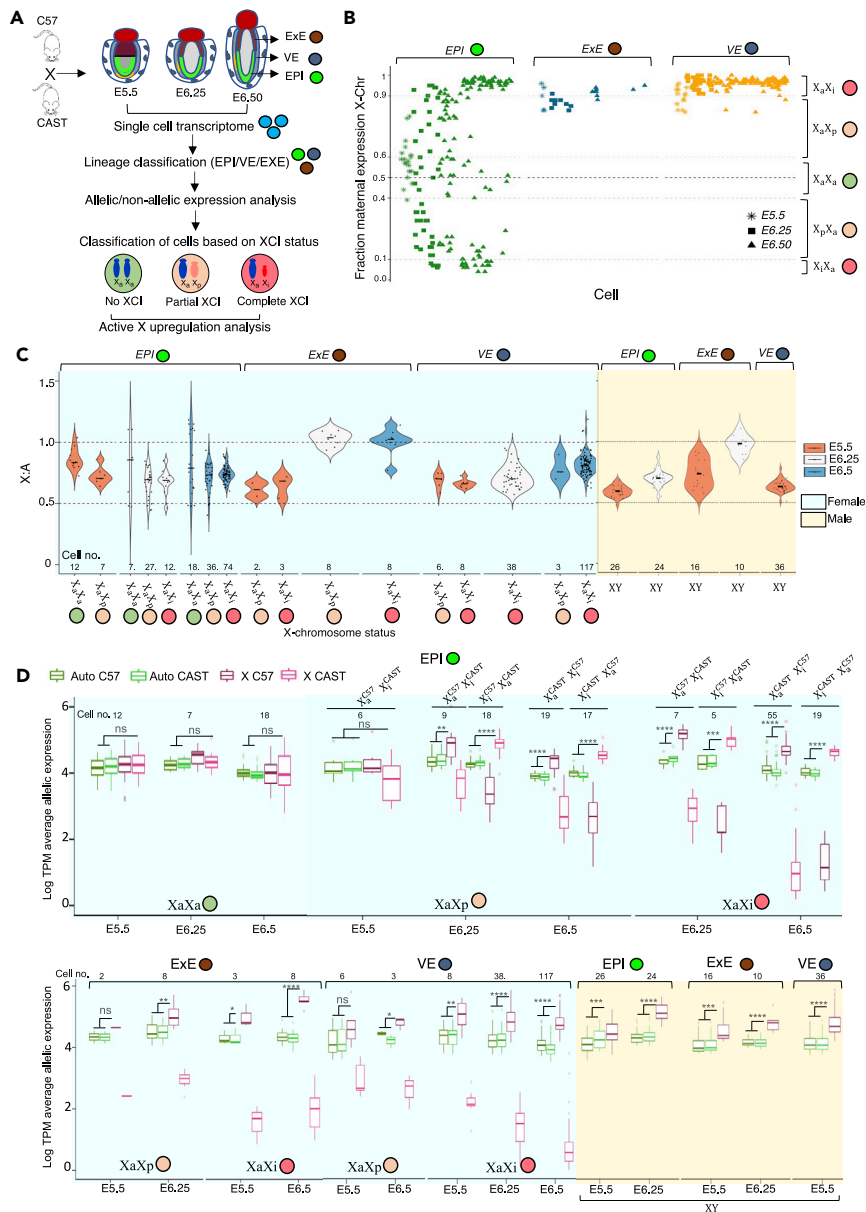
Figure 1. Dynamic X chromosome upregulation in different lineages of pre-gastrulation embryos

(A) Schematic outline of the workflow: profiling of active X-upregulation in different lineages (EPI: Epiblast, ExE: Extra-embryonic ectoderm, and VE: Visceral endoderm) of pre-gastrulation hybrid mouse embryos (E5.5, E6.25, and E6.50) at the single-cell level using scRNA-seq dataset. Hybrid mouse embryos were obtained from crossing between two divergent mouse strains C57 and CAST.

(B) Classification of cells based on XCI state through profiling of fraction maternal expression of X-linked genes in the single cells of different lineages of female pre-gastrulation embryos (EPI, ExE, and VE). Ranges of the fraction of maternal expression were considered for different category cells: X_aX_a = 0.4–0.6, X_aX_p = 0.6–0.9/0.1–0.4 and X_aX_i = 0.9–1/0–0.1.

(C) X:A ratios represented as violin plots for the different lineages of female (XaXa, XaXp, XaXi) and male cells of pre-gastrulation embryos. Line within each violin represents a median value.

(D) Boxplots showing allelic expression levels of X-linked and autosomal genes in different lineages of female (XaXa, XaXp, XaXi) and male cells of pre-gastrulation embryos. Line inside of the each box represents median value, edges of each box indicating 25 and 75% of the dataset. (Mann-Whitney U test: p *value < 0.05; * p value < 0.01; ** and p value < 0.001; ***).
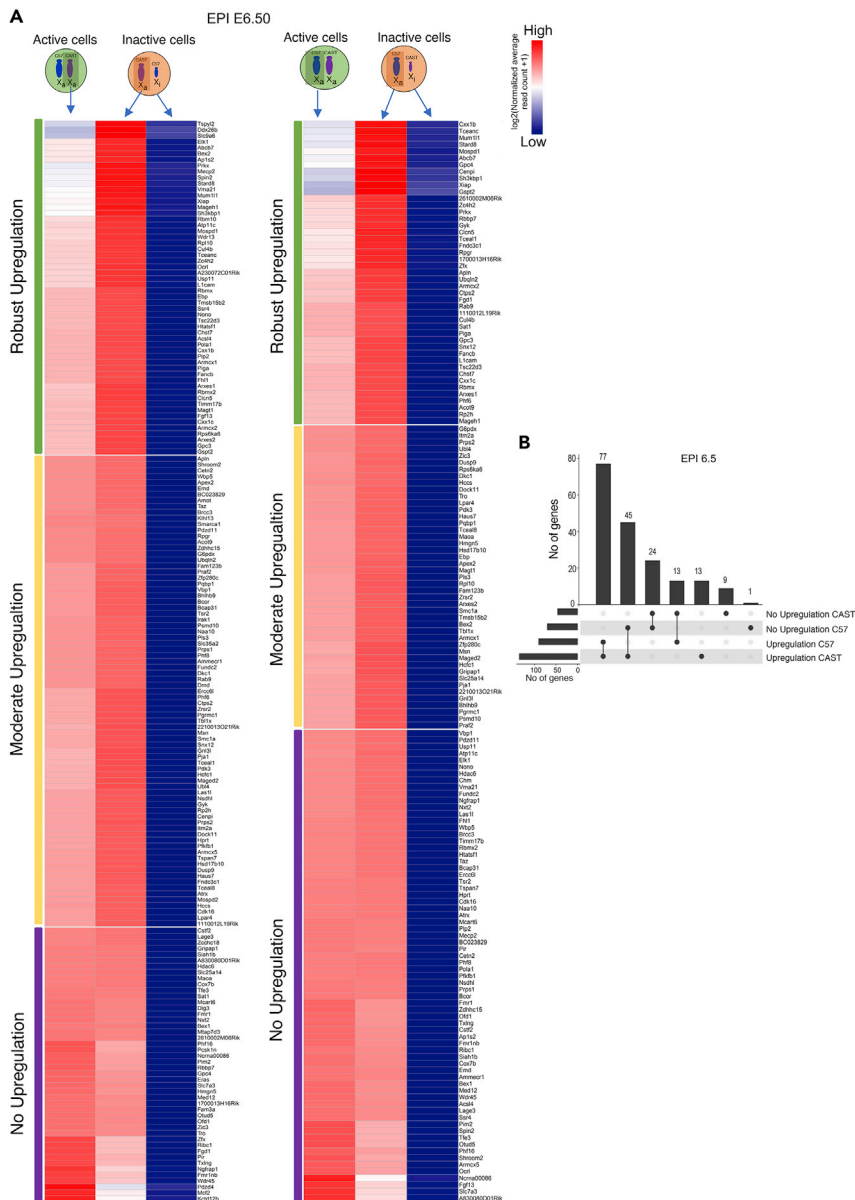
Figure 2. X chromosome upregulation is not global

(A and B) Comparison of expression (log normalized reads) of individual X-linked genes from the active allele between the XaXa and XaXi EPI cells of E6.50 embryos (B) Plot representing intersections of allele-wise upregulation of X-linked genes in E6.5 EPI cells.

Higher occupancy of different activating factors at upregulated X-linked genes versus non-upregulated X-linked genes

Next, we investigated the mechanistic basis of why some genes undergo upregulation and others are not from the same active X. For this purpose, we considered the X-linked genes (in E6.5 EPI) having upregulation from both alleles as upregulated genes, whereas genes showing no upregulation from either of the allele categorized as non-upregulated. Previous reports stated that XCU is restricted only to dosage-sensitive genes such as genes encoding for large protein complexes, transcription factors, proteins involved in signal transduction, and so forth. Therefore, we analyzed if the upregulated genes in EPI E6.5 mostly belong to such dosage-sensitive genes. We identified dosage-sensitive X-linked genes through different gene function databases and haplo-insufficient gene databases as described in STAR Methods. We found that while some upregulated genes fall into dosage-sensitive category, many genes are not (Table S3).
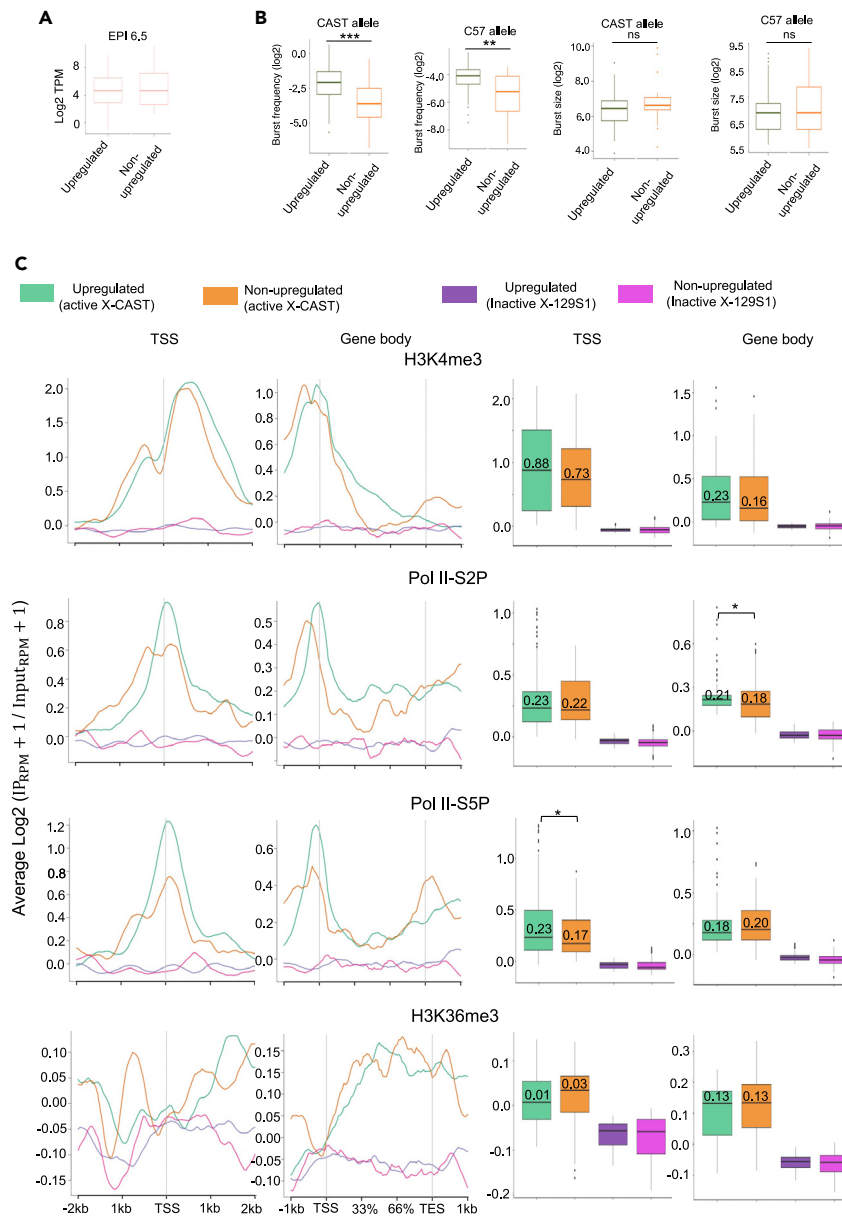
Figure 3. Enrichment of activating factors and increased transcriptional burst frequency leads to the upregulation of X linked genes

(A) Comparison of expression levels (log2 TPM) between upregulated and non-upregulated X-linked genes. In boxplots, the line inside of each box represents median value, edges of each box indicating 25 and 75% of the dataset.

(B) Comparison of allelic burst frequency and burst size of the upregulated vs non-upregulated X-linked genes in the E6.5 EPI cells. In boxplots, the line inside of each box represents median value, edges of each box indicating 25 and 75% of the dataset. (Wilcoxon rank test: p *value < 0.01;* ** and p *value < 0.001;* ***).

(C) Quantitative enrichment analysis of different activating marks (RNA PolII S5P/S2P, H3K4me3, H3K36me3) at TSS and gene body of upregulated and non-upregulated genes in female MEF cells. In boxplots, the line inside of each box represents median value, edges of each box indicating 25 and 75% of the dataset. (Mann-Whitney U test: p *value < 0.05 *).

Moreover, we found that there are many dosage-sensitive X-linked genes among the non-upregulated genes. Altogether, there are no significant differences in the distribution of dosage-sensitive/insensitive genes among the upregulated or non-upregulated genes, suggesting that XCU is not restricted to dosage-sensitive genes only (Table S3). Next, we asked if there were any differences in expression levels between upregulated vs. non-upregulated genes and found no significant differences (Figure 3A). Previous

studies suggested that X-upregulation is associated with enhanced transcriptional burst frequency. To explore this, we profiled the allelic transcriptional burst kinetics between the upregulated and non-upregulated X-linked genes in the X_aX_i cells of E6.5 EPI. Interestingly, the upregulated X-linked genes showed significantly higher burst frequency compared to the not upregulated X-linked genes (Figure 3B). However, burst sizes were not significantly different between the two categories of X-linked genes (Figure 3B). Considering the higher transcriptional burst frequency of the upregulated genes, we next investigated if upregulated genes have higher occupancy of different transcriptional activating factors than non-upregulated genes. To explore this, we profiled the occupancy of RNA PolII-S5P, RNA PolII-S2P, active chromatin marks such as H3K4me3, H3K36me3 at the transcriptional start site (TSS) and gene body of active allele of upregulated and non-upregulated genes loci through allele-specific Chip-Seq analysis of hybrid female mouse embryonic fibroblast (MEF) cells (Figure 3C). These MEF cells harbor skewed inactive X chromosome (129S1) and therefore allowed us to differentiate the enrichment of active marks between active and inactive-X through allele-specific analysis. Indeed, we found that while active allele showed significant enrichment of these different active marks, inactive allele had almost no such enrichment, thus validating our allele-specific Chip-Seq analysis (Figure 3C). Interestingly, we found that the H3K4me3, RNA PolII (S5P/S2P) showed higher occupancy around the TSS, and gene body regions of the upregulated X-linked genes compared to the non-upregulated genes (Figure 3C). However, no upregulated gene-specific enrichment was observed for H3K36me3. Altogether, these data suggested that enhanced occupancy of different activating marks at the upregulated genes loci might lead to higher transcriptional burst frequency, leading to the upregulation of X-linked genes.

In-silico model predicts that recruitment probabilities of activating factors, as well as the availability of activating factors, are keys to the upregulation of X-linked genes

To better understand the mechanisms behind XCU in a quantitative manner, we hypothesized that the difference in response to X-upregulation for different X-linked genes could be caused by a difference in the recruitment rates of activation factors of different genes. To test the validity of the hypothesis, we created an *in silico* model of the X chromosome consisting of two classes of genes, upregulated and non-upregulated. The first category of genes has a relatively higher probability of recruitment of activation factors (see STAR Methods). Based on the simulation, we found that the burst frequency, calculated as the fraction of times a gene turns on, is higher for the upregulated genes compared to non-upregulated genes for XaXi cells (Figure 4A), thus recapitulating our experimental observations in Figure 3B. In contrast, simulation in XaXa cells showed no difference between the two classes of genes (Figure 4A). Additionally, we found that this trend of expression level differences holds across the expression matrices generated using the simulations (Figure 4B). Altogether, this analysis suggested that a difference in the range of recruitment probabilities is sufficient to bring about a difference in mean burst frequencies of upregulated and non-upregulated genes.

We then compared the expression levels of upregulated and non-upregulated genes across XaXa and XaXi cells. The difference in expression levels is statistically significant for the upregulated genes but not so for non-upregulated genes (Figure 4C), suggesting that a difference in recruitment probabilities can also bring about gene upregulation in the active chromosome on XCI. This result is also supported by the fact that within the two kinds of cells (XaXa and XaXi), the difference between upregulated and non-upregulated genes could be seen only for XaXi cells, where there was a difference in the recruitment probabilities (Figures 4D and 4E). We then examined whether the activation factors available to chromosome play any role in determining this difference. We generated expression matrices for different levels of activation factors and found that at lower levels of activation factors, the upregulation is quite noisy, because the downregulated genes also show a significant difference in expression between XaXa and XaXi for more than 60% of cells (Figure 4F). As the number of factors increases, the percentage of downregulated genes showing upregulation decreases, making upregulation less stochastic (Figure 4F).

DISCUSSION

In 1967, Ohno hypothesized that progressive gene loss from the Y chromosome resulted in the upregulation of X chromosomes in both sexes, which subsequently led to the inactivation of one X chromosome in female mammals to ensure proper X-dosage. Although XCI has been studied extensively in recent years, many aspects of XCU remain unexplored. Specially, Ohno's hypothesis is often contested as evidence for active X upregulation remains controversial. Therefore, extensive studies involving different aspects of XCU are crucial.
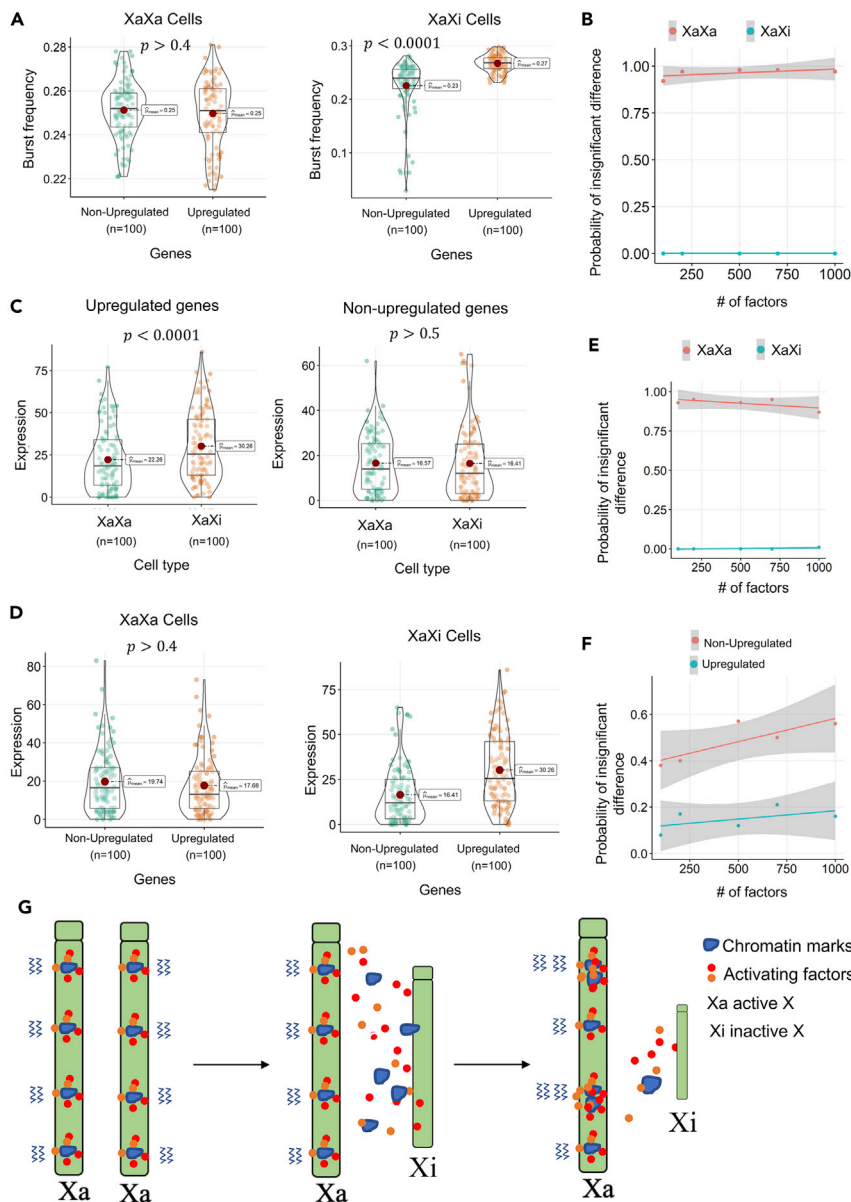
Figure 4. In silico model explains a possible mechanism for XCU

(A) Burst frequency distributions for upregulated and non-upregulated genes in XaXa and XaXi cells from the model. In violin plots included with boxplots, the line inside of the each box represents median value, red dot represents mean, and edges of each box indicating 25 and 75% of the dataset.

(B) Fraction of cells having an insignificant difference between the burst frequency of upregulated and non-upregulated genes for XaXa cells (red), XaXi cells (blue).

(C) Expression level distribution for upregulated and non-upregulated genes from active allele from XaXi (orange) and an allele from XaXa (green) in the model. In violin plots included with boxplots, the line inside of the each box represents median value, red dot represents mean and edges of each box indicating 25 and 75% of the dataset.

(D) Difference in the expression levels of upregulated vs non-upregulated genes of XaXa and XaXi cells. In violin plots included with boxplots, the line inside of the each box represents median value, red dot represents mean and edges of each box indicating 25 and 75% of the dataset.

(E) Same as B but for difference in expression levels of up and non-upregulated genes.

(F) Fraction of in-silico cells that show an insignificant difference in the expression levels of upregulated genes of XaXa and XaXi (red) and non-upregulated genes between XaXa and XaXi (blue).

(G) Model representing that increase of activating factors availability upon XCI and followed by enrichment of these activating factors at the loci of genes with higher recruitment probability on the active X chromosome leads to the upregulation of those X-linked genes.

Here, we have profiled the XCU dynamics at the single-cell level in pre-gastrulation mouse embryos. We found that in EPI cells (E5.5, E6.25, E6.5), which is at the onset of random XCI, XCU is dynamically linked with the XCI. Moreover, VE and ExE cells, which undergo imprinted XCI, also showed the presence of up-regulated active-X chromosome. Altogether, our results extend support for Ohno's hypothesis and suggest that two X chromosomes expression state is highly plastic toward balancing the optimal X chromosome dosage during pre-gastrulation. Our result is consistent with a recent report by Lentini et al. (2021).

Interestingly, analysis of gene-wise dynamics of XCU revealed that many X-linked genes do not undergo upregulation (Figure 2A). This was also reflected in the X:A ratio analysis, where we found that X:A ratio of XaXi cells tends to be lower compared to the XaXa cells (Figure 1C). Therefore, we suggested that though XCU is dynamically linked to XCI, XCU is might not global or chromosome-wide like XCI. However, one caveat of our analysis is that we are unable to profile gene-wise dynamics along the X chromosome from these scRNA-seq datasets as allelic expression of many X-linked genes was unknown. One of the major limitations of scRNA-seq is low capture efficiency and high dropouts owing to the very low starting material (Chen et al., 2019). Moreover, to ensure the quality of our allelic analysis of gene expression, we applied many filters, which led to exclusion of many X-linked genes from our analysis. Therefore, in the future, more extensive studies are necessary to profile gene-wise dynamics along the X chromosome with many more X-linked genes to better understand the non-uniform nature of XCU.

On the other hand, previous studies implicated that XCU affects dosage-sensitive genes such as components of macromolecular complexes, signal transduction pathways, or encoding for transcription factors (Pessia et al., 2012, 2014). However, we found that upregulated genes are not restricted to dosage-sensitive genes only and there are many dosage-sensitive genes that do not undergo upregulation (Table S3). Another possibility has been proposed that dosage compensation can be achieved through downregulating autosomal genes, which are interacting with dosage-sensitive X-linked genes (Julien et al., 2012). In that context, it would be interesting to explore if there is a link between recently described random monoallelically expressed autosomal genes with the X-linked genes. On the other hand, depletion of haplo-insufficient genes from the X chromosome has been reported and which can lead to the avoiding of the need for global XCU (de Clare et al., 2011). Another way could have happened that haplo-insufficient genes transposed to the autosomes and thereby they can express from two copies, which can also help in avoiding global XCU.

Separately, it is believed that the evolution of XCI happened to counteract the upregulation of X in female cells, as proposed by Ohno. If so, we wondered why XCI is global and occurs in the majority of X-linked genes while XCU is not. It may be possible that XCI initially evolved from the genes undergoing upregulation and later became a global phenomenon for some unknown reasons. Indeed, it is thought that XCI evolved in region-specific manner started from where Y degradation started first and gradually spread over the other degraded regions as indicated by regional differences of several epigenetic modifications involved in XCI (Chadwick and Willard, 2004; Prothero et al., 2009). Subsequently, the gradual accumulation of different sequences on the X chromosome might have played a booster role in the propagation of XCI signal chromosome-wide (Jegalian and Page, 1998; Lyon, 1998; Prothero et al., 2009). On the other hand, other possibilities have been proposed that consider XCI evolved first and subsequently, XCU came into the picture. Haig proposed an alternative model (parental antagonism model) of the evolution of XCI, suggesting that XCI may have evolved initially in the form of genomic imprinting related to parental conflicts (Engelstädter and Haig, 2008; Haig, 2006; Pessia et al., 2014). Haig's hypothesis also predicts that imprinted paternal XCI rather than the random XCI is the ancient form of XCI. Indeed, in marsupials and some tissues of placentals, XCI always occurs on paternal X (Cooper et al., 1993; Harris et al., 2019; Maclary et al., 2014, 2017). However, further investigation is necessary to validate Haig's model of the evolution of XCI or to test whether XCI evolved from a completely different perspective.

Next, our analysis unveils the mechanistic basis of how some genes get upregulated and others are not. Previous studies have reported that the X-linked genes from active X chromosome have more significant enrichment of the active chromatin marks compared to the autosomal genes (Deng et al., 2013). We found that upregulated X-linked genes have a higher occupancy of different transcriptional activating marks such as H3K4me3, RNA PolII S2P (gene body), RNA PolII S5P (TSS) compared to the non-upregulated genes (Figure 3C). Interestingly, transcriptional burst frequencies were significantly enhanced in upregulated genes than non-upregulated genes (Figure 3B). Taken together, our results unveil that the enrichment of
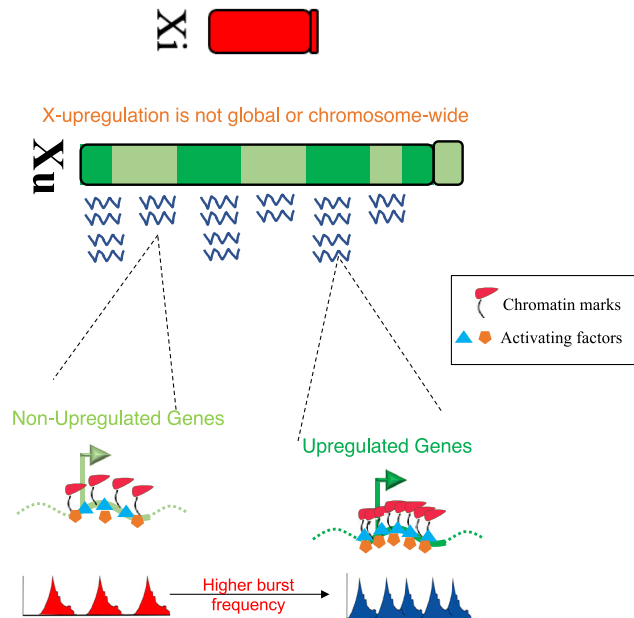
Figure 5. Model representing X-upregulation dynamics and mechanisms:
X-upregulation is not global or chromosome-wide. Upregulated X-linked genes show higher occupancy of different active marks and have higher transcriptional burst frequency compared to the non-upregulated genes.

activating marks might lead to higher transcriptional burst frequencies and thereby results in the upregulation of X-linked genes. Additionally, our in-silico analysis shows that a difference in the range of recruitment probabilities of different activating factors is sufficient to bring about a difference in the mean burst frequencies of upregulated and non-upregulated genes (Figure 4). Importantly, the availability of the activation factors also plays an important role in determining these differences. We predict that on XCI, numerous *trans*-acting factors leave the inactivating-X leading to a global increase in the number of activating factors. We show that on such an increase of activating factors availability, genes with higher recruitment probability get enriched with these factors and bring the upregulation of those genes (Figures 4G and 5). Interestingly, although upregulated genes have increased the ability to recruit activators, yet do not exhibit an overall higher expression level compared to the non-upregulated genes, suggesting that activators enrichment does not proportionate with the transcriptional output (Figure 3A). Indeed, previous studies have reported a nonlinear relationship between chromatin marks and transcriptional output (Yildirim et al., 2012).

Limitation of the study

Here, we have profiled gene-wise dynamics of XCU in pre-gastrulation mouse embryos at the single-cell level through the allele-specific analysis of scRNA-seq dataset. One limitation of our study is that though we have profiled XCU dynamics for the substantial number of X-linked genes, many X-linked genes were excluded from our analysis to ensure the quality of our scRNAseq analysis. Therefore, in the future, more extensive studies are necessary to get better insight into the gene-expression dynamics along the X chromosome.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104465.

AUTHOR CONTRIBUTIONS

SG conceptualized and supervised the study. HCN and DC performed data analyses. KH and MKJ worked on mathematical modeling. SG, MKJ, HCN, and KH wrote, edited, and proofread the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638.

Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Ruepp, A. (2019). CORUM: the comprehensive resource of mammalian protein complexes-2019. Nucleic Acids Res. *47*, D559–D563. https://doi.org/10.1093/NAR/GKY973.

Hu, H., Miao, Y.R., Jia, L.H., Yu, Q.Y., Zhang, Q., and Guo, A.Y. (2019). AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. Nucleic Acids Res. *47*, D33–D38. https://doi.org/10.1093/NAR/GKY822.

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P.D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and im-provements in enrichment analysis tools. Nucleic Acids Res. *47*, D419–D426. https://doi.org/10.1093/NAR/GKY1038.

Pacini, G., Dunkel, I., Mages, N., Mutzel, V., Timmermann, B., Marsico, A., and Schulz, E.G. (2021). Integrated analysis of Xist upregulation and X-chromosome inactivation with single-cell and single-allele resolution. Nat. Commun. *12*, 3638. https://doi.org/10.1038/S41467-021-23643-6.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

Chadwick, B.P., and Willard, H.F. (2004). Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. Proc. Natl. Acad. Sci. U S A *101*, 17450–17455. https://doi.org/10.1073/PNAS.0408021101.

Chen, X., and Zhang, J. (2016). The X to autosome expression ratio in haploid and diploid human embryonic stem cells. Mol. Biol. Evol. *33*, 3104–3107. https://doi.org/10.1093/molbev/msw187.

Chen, G., Ning, B., and Shi, T. (2019). Single-cell RNA-seq technologies and related computational data analysis. Front. Genet. *10*, 317. https://doi.org/10.3389/FGENE.2019.00317/BIBTEX.

Cheng, S., Pei, Y., He, L., Peng, G., Reinius, B., Tam, P.P.L., Jing, N., and Deng, Q. (2019). Single-cell RNA-seq reveals cellular heterogeneity of pluripotency transition and X chromosome dynamics during early mouse development. Cell Rep. *26*, 2593–2607.e3. https://doi.org/10.1016/j.celrep.2019.02.031.

Cooper, D.W., Johnston, P.G., Watson, J.M., and Graves, J.A.M. (1993). X-inactivation in marsupials and monotremes. Semin. Dev. Biol. *4*, 117–128. https://doi.org/10.1006/SEDB.1993.1014.

Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grützner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. Nature *508*, 488–493. https://doi.org/10.1038/nature13151.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

de Clare, M., Pir, P., and Oliver, S.G. (2011). Haploinsufficiency and the sex chromosomes from yeasts to humans. BMC Biol. 9, 15. https://doi.org/10.1186/1741-7007-9-15/TABLES/4.

Moreira de Mello, J.C., Fernandes, G.R., Vibranovski, M.D., and Pereira, L.V. (2017). Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. Sci. Rep. 7, 10794. https://doi.org/10.1038/s41598-017-11044-z.

Deng, X., Hiatt, J.B., Nguyen, D.K., Ercan, S., Sturgill, D., Hillier, L.W., Schlesinger, F., Davis, C.A., Reinke, V.J., Gingeras, T.R., et al. (2011). Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. Nat. Genet. 43, 1179–1185. https://doi.org/10.1038/ng.948.

Deng, X., Berletch, J.B., Ma, W., Nguyen, D.K., Hiatt, J.B., Noble, W.S., Shendure, J., and Disteche, C.M. (2013). Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. Dev. Cell 25, 55–68. https://doi.org/10.1016/j.devcel.2013.01.028.

Nguyen, D.K., and Disteche, C.M. (2006). Dosage compensation of the active X chromosome in mammals. Nat. Genet. 38, 47–53. https://doi.org/10.1038/ng1705.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Engelstädter, J., and Haig, D. (2008). Sexual antagonism and the evolution of X chromosome inactivation. Evolution 62, 2097–2104. https://doi.org/10.1111/J.1558-5646.2008.00431.X.

Gayen, S., Maclary, E., Buttigieg, E., Hinten, M., and Kalantry, S. (2015). A primary role for the tsix lncRNA in maintaining random X-chromosome inactivation. Cell Rep. 11, 1251–1265. https://doi.org/10.1016/j.celrep.2015.04.039.

Gayen, S., Maclary, E., Hinten, M., and Kalantry, S. (2016). Sex-specific silencing of X-linked genes by Xist RNA. Proc. Natl. Acad. Sci. U S A 113. https://doi.org/10.1073/pnas.1515971113.

Graves, J.A.M. (2016). Evolution of vertebrate sex chromosomes and dosage compensation. Nat. Rev. Genet. 17, 33–46. https://doi.org/10.1038/nrg.2015.2.

Gupta, V., Parisi, M., Sturgill, D., Nuttall, R., Doctolero, M., Dudko, O.K., Malley, J.D., Eastman, P.S., and Oliver, B. (2006). Global analysis of X-chromosome dosage compensation. J. Biol. 5, 3. https://doi.org/10.1186/jbiol30.

Haig, D. (2006). Self-imposed silence: parental antagonism and the evolution of X-chromosome inactivation. Evolution 60, 440–447. https://doi.org/10.1111/J.0014-3820.2006.TB01125.X.

Harris, C., Cloutier, M., Trotter, M., Hinten, M., Gayen, S., Du, Z., Xie, W., and Kalantry, S. (2019). Conversion of random x-inactivation to imprinted x-inactivation by maternal prc2. Elife 8. https://doi.org/10.7554/eLife.44258.

Rehm, H.L., Berg, J.S., Brooks, L.D., Bustamante, C.D., Evans, J.P., Landrum, M.J., Ledbetter, D.H., Maglott, D.R., Martin, C.L., Nussbaum, R.L., et al. (2015). ClinGen–the clinical genome resource. N. Engl. J. Med. 372, 2235–2242. https://doi.org/10.1056/NEJMSR1406261.

Hounkpe, B.W., Chenou, F., de Lima, F., and De Paula, E.V. (2021). HRT Atlas v1.0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. Nucleic Acids Res. 49, D947–D955. https://doi.org/10.1093/NAR/GKAA609.

Jegalian, K., and Page, D.C. (1998). A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature 394, 776–780. https://doi.org/10.1038/29522.

Jiang, Y., Zhang, N.R., and Li, M. (2017). SCALE: modeling allele-specific gene expression by single-cell RNA sequencing. Genome Biol. 18, 74. https://doi.org/10.1186/s13059-017-1200-8.

Johnston, C.M., Lovell, F.L., Leongamornlert, D.A., Stranger, B.E., Dermitzakis, E.T., and Ross, M.T. (2008). Large-scale population study of human cell lines indicates that dosage compensation is virtually complete. PLoS Genet. 4, e9. https://doi.org/10.1371/journal.pgen.0040009.

Julien, P., Brawand, D., Soumillon, M., Necsulea, A., Liechti, A., Schütz, F., Daish, T., Grützner, F., and Kaessmann, H. (2012). Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. PLoS Biol. 10, e1001328. https://doi.org/10.1371/journal.pbio.1001328.

Kharchenko, P.V., Xi, R., and Park, P.J. (2011). Evidence for dosage compensation between the X chromosome and autosomes in mammals. Nat. Genet. 43, 1167–1169. https://doi.org/10.1038/ng.991.

Krueger, F., and Andrews, S.R. (2016). SNPsplit: allele-specific splitting of alignments between genomes with known SNP genotypes. F1000Research 5, 1479. https://doi.org/10.12688/F1000RESEARCH.9037.2.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/NMETH.1923.

Larsson, A.J.M., Coucoravas, C., Sandberg, R., and Reinius, B. (2019). X-chromosome upregulation is driven by increased burst frequency. Nat. Struct. Mol. Biol. 26, 963–969. https://doi.org/10.1038/s41594-019-0306-y.

Lentini, A., Cheng, H., Noble, J., Papanicolaou, N., Coucoravas, C., Andrews, N., Deng, Q., Enge, M., and Reinius, B. (2021). Allele-resolved single-cell multi-omics uncovers the dynamics and transcriptional kinetics of x-chromosome upregulation. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.14.452323.

Li, W.V., and Li, J.J. (2018). An accurate and robust imputation method scImpute for single-cell RNA-seq data. Nat. Commun. 9, 997. https://doi.org/10.1038/s41467-018-03405-7.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017a). A comprehensive mouse transcriptomic BodyMap across 17 tissues by RNA-seq. Sci. Rep. 7, 4200. https://doi.org/10.1038/s41598-017-04520-z.

Li, X., Hu, Z., Yu, X., Zhang, C., Ma, B., He, L., Wei, C., and Wu, J. (2017b). Dosage compensation in the process of inactivation/reactivation during both germ cell development and early embryogenesis in mouse. Sci. Rep. 7, 3729. https://doi.org/10.1038/s41598-017-03829-z.

Lin, H., Gupta, V., Vermilyea, M.D., Falciani, F., Lee, J.T., O'Neill, L.P., and Turner, B.M. (2007). Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. PLoS Biol. 5, e326. https://doi.org/10.1371/journal.pbio.0050326.

Lin, H., Halsall, J.A., Antczak, P., O'Neill, L.P., Falciani, F., and Turner, B.M. (2011). Relative overexpression of X-linked genes in mouse embryonic stem cells is consistent with Ohno's hypothesis. Nat. Genet. 43, 1169–1170. https://doi.org/10.1038/ng.992.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (mus musculus L.). Nature 190, 372–373. https://doi.org/10.1038/190372a0.

Lyon, M.F. (1998). X-Chromosome inactivation: a repeat hypothesis. Cytogenet. Genome Res. 80, 133–137. https://doi.org/10.1159/000014969.

Maclary, E., Buttigieg, E., Hinten, M., Gayen, S., Harris, C., Sarkar, M.K., Purushothaman, S., and Kalantry, S. (2014). Differentiation-dependent requirement of Tsix long non-coding RNA in imprinted X-chromosome inactivation. Nat. Commun. 5, 4209. https://doi.org/10.1038/ncomms5209.

Maclary, E., Hinten, M., Harris, C., Sethuraman, S., Gayen, S., and Kalantry, S. (2017). PRC2 represses transcribed genes on the imprinted inactive X chromosome in mice. Genome Biol. 18, 82. https://doi.org/10.1186/s13059-017-1211-5.

Mandal, S., Chandel, D., Kaur, H., Majumdar, S., Arava, M., and 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. https://doi.org/10.1016/j.stemcr.2020.03.027.

Naik, H.C., Hari, K., Chandel, D., Mandal, S., Jolly, M.K., and Gayen, S. (2021). Semicoordinated allelic-bursting shape dynamic random monoallelic expression in pregastrulation embryos. iScience 24, 102954. https://doi.org/10.1016/J.ISCI.2021.102954.

Ohno, S. (1967). Sex Chromosomes and Sex-Linked Genes, *vol 68* (Springer-Verlag), p. 1375. https://doi.org/10.7326/0003-4819-68-6-1375_2.

Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., and Marais, G.A.B. (2012). Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proc. Natl. Acad. Sci. U S A *109*, 5346–5351. https://doi.org/10.1073/pnas.1116763109.

Pessia, E., Engelstädter, J., and Marais, G.A.B. (2014). The evolution of X chromosome inactivation in mammals: the demise of Ohno's hypothesis? Cell. Mol. Life Sci. *71*, 1383–1394. https://doi.org/10.1007/s00018-013-1499-6.

Prothero, K.E., Stahl, J.M., and Carrel, L. (2009). Dosage compensation and gene expression on the mammalian X chromosome: one plus one does not always equal two. Chromosome Res. *17*, 637–648. https://doi.org/10.1007/S10577-009-9063-9.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Sangrithi, M.N., Royo, H., Mahadevaiah, S.K., Ojarikre, O., Bhaw, L., Sesay, A., Peters, A.H.F.M., Stadler, M., and Turner, J.M.A. (2017). Non-canonical and sexually Dimorphic X dosage compensation states in the mouse and human germline. Dev. Cell *40*, 289–301.e3. https://doi.org/10.1016/j.devcel.2016.12.023.

Sarkar, M.K., Gayen, S., Kumar, S., Maclary, E., Buttigieg, E., Hinten, M., Kumari, A., Harris, C., Sado, T., and Kalantry, S. (2015). An Xist-activating antisense RNA required for X-chromosome inactivation. Nat. Commun. *6*, 8564. https://doi.org/10.1038/ncomms9564.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). Ngs.plot: quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genom. *15*, 284. https://doi.org/10.1186/1471-2164-15-284/FIGURES/6.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

Sun, K., Wang, H., and Sun, H. (2017). mTFkb: a knowledgebase for fundamental annotation of mouse transcription factors. Sci. Rep. *7*, 3022. https://doi.org/10.1038/s41598-017-02404-w.

Talebizadeh, Z., Simon, S.D., and Butler, M.G. (2006). X chromosome gene expression in human tissues: male and female comparisons. Genomics *88*, 675–681. https://doi.org/10.1016/j.ygeno.2006.07.016.

Veyrunes, F., Waters, P.D., Miethke, P., Rens, W., McMillan, D., Alsop, A.E., Grutzner, F., Deakin, J.E., Whittington, C.M., Schatzkamer, K., et al. (2008). Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res. *18*, 965–973. https://doi.org/10.1101/gr.7101908.

Xiong, Y., Chen, X., Chen, Z., Wang, X., Shi, S., Wang, X., Zhang, J., and He, X. (2010). RNA sequencing shows no dosage compensation of the active X-chromosome. Nat. Genet. *42*, 1043–1047. https://doi.org/10.1038/ng.711.

Yildirim, E., Sadreyev, R.I., Pinter, S.F., and Lee, J.T. (2012). X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. Nat. Struct. Mol. Biol. *19*, 56–61. https://doi.org/10.1038/nsmb.2195.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Seurat (v3.1.5)	(Butler et al., 2018; Stuart et al., 2019)	https://satijalab.org/seurat/
STAR (v2.5.4b)	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
scImpute (v0.0.9)	(Li and Li, 2018)	https://github.com/Vivianstats/scImpute
HTSeq-count (v0.13.5)	(Anders et al., 2015)	https://htseq.readthedocs.io/en/master/index.html
SCALE (v1.3.0)	(Jiang et al., 2017)	https://github.com/yuchaojiang/SCALE
Bowtie2 (v2.3.4.1)	(Langmead and Salzberg, 2012)	https://github.com/BenLangmead/bowtie2
SNPsplit (v0.5.0)	(Krueger and Andrews, 2016)	https://github.com/FelixKrueger/SNPsplit
Samtools (v1.7)	(Li et al., 2009)	http://www.htslib.org/
BEDTools (v2.26.0)	(Quinlan and Hall, 2010)	https://github.com/arq5x/bedtools2
VCF tools (v0.1.17)	(Danecek et al., 2011)	https://github.com/vcftools/vcftools
ngs.plot (v2.61)	(Shen et al., 2014)	https://github.com/shenlab-sinai/ngsplot
R (v4.1.3)	R core team	https://www.R-project.org/
Code for simulation	This paper	https://github.com/csbBSSE/XCU

RESOURCE AVAILABILITY

Lead contact

Further information on resources and reagents should be directed to and will be fulfilled by the Lead Contact, Srimonta Gayen (srimonta@iisc.ac.in).

Materials availability

This study did not generate new unique reagents.

Data and code availability

This study did not generate any unique datasets. The code used for simulation is available at https://github.com/csbBSSE/XCU For questions regarding the raw data from the current study, please contact the lead contact. All software's used in this study are commercially available.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

No experimental model system used for this study.

METHOD DETAILS

Data acquisition

Pre-processed single-cell RNA-seq dataset for pre-gastrulation embryos was retrieved from GSE109071 (Cheng et al., 2019). In this dataset, E5.5 and E6.25 embryos were derived from hybrid mouse embryos (C57BL/6J:female × CAST/EiJ:male), whereas E6.5 embryos were derived from reciprocal cross (CAST/EiJ:female × C57BL/6J:male). ChIP-seq datasets were obtained from GSE33823 (Yildirim et al., 2012), for H3K4me3, H3K36me3, RNAPolII S5P/S2P.

Read alignment and counting

RNA-seq reads were aligned to the mouse reference genome mm10 using STAR (Dobin et al., 2013) and counted using HTSeq-count (Anders et al., 2015). To avoid the dropout events due to low amount of starting material or failed amplification of original RNAs in scRNA-Seq, we used a statistical imputation method scImpute (Li and Li, 2018). scImpute identifies the likely dropouts and do imputation without introducing any bias in the rest of the data. We used scImpute R package (v0.0.9) with parameter Kcluster = 3, but

otherwise, default parameters. Expression levels of transcripts was computed using *Transcripts per million* (TPM) method. We performed Kolmogorov-Smirnov's test to identify cells (9-cells) with significant differences in the gene expression distribution between X and autosomal genes and removed from downstream analysis.

Sexing of the embryos

For sex-determination of the pre-gastrulation embryos, an embryo was classified as male if the sum of the read count for the Y-linked genes (Usp9y, Uty, Ddx3y, Eif2s3y, Kdm5d, Ube1y1, Zfy2, Zfy1) in each cell of an embryo was greater than 12, rest were considered as female embryos.

Lineage identification

We collected the data related to classification of different lineages (EPI, ExE and VE) of pre-gastrulation embryos from our previous work (Naik et al., 2021). In brief, all single cells were subjected to t-distributed stochastic neighbour embedding (t-SNE) to cluster cells and identify lineages. t-SNE was performed using Seurat (version 3.1.5) (Butler et al., 2018; Stuart et al., 2019). Three thousand most variable genes were used for the analysis. Shared Nearest Neighbor (SNN) modularity optimization-based clustering algorithm 'Findcluster' were used to identify cell clusters. Each cluster was allocated to cell lineages based on the expression of specific marker genes: Oct4 (EPI), Bmp4 (ExE), and Amn (VE).

X:A ratio analysis

In view of large number of autosomal genes compare to the small set of X-linked genes (14442-15497 autosomal genes, 798-733 X-linked genes), we calculated the X:A ratio for different lineages of pre-gastrulation embryo through bootstrapping procedure adapted from previous studies (Pacini et al., 2021). For each cell, we divided expression of X-linked genes with same number of randomly selected autosomal genes and this step was repeated 1000 times and median of this 1000 values were considered for X/A expression ratio. For this analysis, we used those X-linked and autosomal genes, which had ≥ 0.5 TPM and minimum 10 cell expression in each respective lineage. We excluded the escapees of X-inactivation and the genes in the pseudo autosomal region from our analysis.

Allele-specific expression analysis

We first constructed in silico CAST specific parental genome by incorporating CAST/EiJ specific SNPs (https://www.sanger.ac.uk/science/data/mouse-genomes-project) into the mm10 genome using VCF tools (Danecek et al., 2011; Li et al., 2009; Quinlan and Hall, 2010). Reads were mapped onto C57BL/6J (mm10) reference genome and in silico CAST genomes using STAR allowing no multi-mapped reads. To exclude any false positives, we only considered those genes with at least 2 informative SNPs and minimum 3 reads per SNP site. We took an average of SNP-wise reads to have the allelic read counts. We normalized allelic read counts using Spike-in control as described in Naik et al. (2021) (Naik et al., 2021). In brief, we calculated the sum of reads mapping to all Spike-in molecule in individual cell and then we divided each cell's Spike-in reads with the highest Spike in value to deduce the normalization factors. Finally, allelic counts of individual cells were normalized dividing by the corresponding normalization factor. After normalization of allelic read counts, we considered only those genes for downstream analysis which had at least 1 average reads across the cells of each lineage from a specific developmental stage for pre-gastrulation embryos. Further, only those single-cells were considered for downstream analysis which showed at least 10 X-linked gene expressions. Allelic ratio was obtained individually for each gene using formula = (Maternal/Paternal reads) ÷ (Maternal reads + Paternal reads).

Transcriptional burst kinetics

We used SCALE to profile allelic transcriptional burst kinetics (Jiang et al., 2017). For this analysis, we used genes with at least 5 average read counts across the cells of EPI E6.5 considering that single cells are more prone to dropout event. Additionally, escapee genes were removed from our analysis.

ChIP-seq analysis

First, we created An 'N-masked genome in-silico using SNPsplit (0.4.0) (Krueger and Andrews, 2016). The reads for all samples are then mapped to this N-masked genome with Bowtie2 (-N 1) (Langmead and Salzberg, 2012). Duplicate reads were removed with Picard 'MarkDuplicates' ['REMOVE_DUPLICATES = TRUE'] and blacklisted regions were removed according to encode consortium. SNPsplit was then used

to create allele-specific BAM files by segregating the aligned reads into two distinct alleles (129S1/SvImJ and CAST/EiJ). Enrichment plots and quantification were created Using ngs.plot (-AL bin -MW 15) (Shen et al., 2014).

Identification of dosage sensitive genes

To identify dosage sensitive X-linked genes, we profiled the associated biological function of the genes through Signal related function (http://geneontology.org/) (Mi et al., 2019), Transcription factor databases (http://bioinfo.life.hust.edu.cn/AnimalTFDB/#!/, https://sunlab.cpy.cuhk.edu.hk/mTFkb/) (Hu et al., 2019; Sun et al., 2017), genes involved in protein complex databases (http://mips.helmholtz-muenchen.de/corum/#) (Giurgiu et al., 2019), dosage sensitive genes (human) database (https://www.clinicalgenome.org/) (Rehm et al., 2015), Housekeeping genes databases (https://housekeeping.unicamp.br/) (Li et al., 2017a, 2017b) (Hounkpe et al., 2021; Li et al., 2017a). We called a gene as dosage sensitive if it has association with at least one of the category/databases.

Simulation

The *in-silico* chromosome: The chromosome is modeled as a set of genes. Each gene has a probability of recruitment of transcriptional activation factors (epigenetic factors, transcription factors etc.). Based on probability, genes were classified into two categories: upregulated and non-upregulated, with decreasing probabilities respectively. We took 100 each of these genes and positioned them at random positions on the chromosome. The recruitment happens from a column of activation factors, with 1000 molecules placed in front of every gene. At every time step, each gene can recruit an activation factor from the corresponding position on the column, depending upon the recruitment probability of the gene. The activation factors are modeled to have a discrete diffusion, in that the reduction in activation factors at any position is filled in by the neighboring positions. The recruitment onto the gene is modeled as a cooperative process, with the probability of recruitment increasing sigmoidally with increase in the extent of recruitment. Each gene has a probability of activation, also modeled as a sigmoidal function of the extent of recruitment on the gene. The probability of inactivation for each gene is 0.5. An active gene produces one RNA molecule at every time step. These RNA molecules degrade with a probability of 0.35.

Simulation procedure: At each time step, the recruitment of activation factors happens to all genes if the recruitment probability for the genes is higher than a uniformly random number generated. After recruitment, diffusion is carried out to normalize the activation factor column. Then, genes are turned on/off with the on/off probability as described above. All active genes produce one RNA molecule (making the timescale of the simulation the same as transcription). RNA degradation also is implemented one molecule at a time with a probability of 0.35. For sensitivity analysis, we generated 100 *in silico* cells, each represented by a chromosome described above.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed using the R software (https://www.R-project.org/). Mann-Whitney U two-sided test was used for statistical significance analysis and p values < 0.05 was considered as significant.