

Phosphate Starvation Drives Directional Repositioning of PHO Loci to the Nuclear Periphery and Amplifies Transcriptional Bursting in *Arabidopsis thaliana*

Sudipta Joardar^{1,2}, Swapnil Anil Prabhudesai^{3,4}

1. *Towards Bioscience* (towardsbioscience.com), India.

2. *University of Bologna, Bologna, Italy.*

3. Department of Basic Sciences & Humanities, SVKM's NMIMS Mukesh Patel School of Technology Management & Engineering, Mumbai, India

4. Heinz Maier-Leibnitz Zentrum (MLZ), Technische Universität München, Lichtenbergstraße 1, 85748 Garching, Germany

Correspondence: joardars2025@gmail.com | sudipta.joardar@studio.unibo.it

Abstract

The spatial organization of gene loci within the nucleus profoundly influences transcriptional output. Phosphate (Pi) starvation triggers a rapid and sweeping transcriptional reprogramming in *Arabidopsis thaliana*, yet whether this involves physical repositioning of Pi-responsive loci (PHO genes) within the nuclear volume has not been investigated. Here, we introduce the Nuclear Periphery Burst Index (NPBI), a new composite metric that integrates 3D genomic position with transcriptional burst frequency to quantify the functional coupling between peripheral nuclear localization and stochastic transcriptional activation. Using integrative computational modeling combining polymer-physics-based 3D genome reconstruction from Hi-C contact data, telegraphic burst kinetics simulation at single-locus resolution, topological data analysis (TDA) via graph Laplacian eigenspectrum, and spatial logistic regression, we demonstrate that PHO gene loci undergo directional repositioning toward the nuclear periphery within 24 hours of Pi withdrawal. This repositioning is accompanied by a 1.8-fold increase in NPBI ($p = 0.0002$, Mann-Whitney U), indicating a strong spatial coupling between peripheral localization and elevated burst frequency. Topological analysis reveals a 39% increase in spectral gap and higher persistence entropy in the -Pi genome, suggesting increased long-range chromatin connectivity. Spatial logistic regression confirms that normalized radial position is a significant predictor of high-burst status ($\beta = 2.31$). Together, these findings establish peripheral repositioning as a component of the Pi-starvation transcriptional response and introduce NPBI as a generalizable metric for studying spatial regulation of transcriptional noise in plants.

Keywords: *nuclear architecture; phosphate starvation; transcriptional bursting; nuclear periphery; 3D genome; topological data analysis; Arabidopsis thaliana; Nuclear Periphery Burst Index; chromatin repositioning; stochastic gene expression*

1. Introduction

The spatial organization of chromosomes within the eukaryotic nucleus is not a passive consequence of genome packaging but an active regulatory layer that shapes transcriptional output, DNA repair fidelity, and epigenetic memory (Dekker & Mirny, 2016; Rowley & Corces, 2018). At the submicron scale, individual gene loci occupy non-random positions within the nuclear volume, and evidence from animal cells has established that proximity to specific nuclear compartments — including the nuclear lamina, nuclear pore complexes (NPCs), and the nucleolus — is associated with defined transcriptional states (van Steensel & Belmont, 2017). Genes tethered to the nuclear lamina reside in repressive lamina-associated domains (LADs), while genes repositioned to the nuclear periphery near NPCs are often transcriptionally activated through the "gene gating" mechanism (Blobel, 1985; Casolari et al., 2004).

In plants, nuclear architecture has emerged as a critical integrator of environmental signals (Joardar, 2026). The nuclear periphery is populated by CROWDED NUCLEI (CRWN) proteins that form a functional analog of the animal nuclear lamina, anchoring heterochromatic chromocenters and maintaining chromosome organization in the Rabl configuration (Wang et al., 2013). Light, temperature, mechanical stress, and nutrient signals all influence 3D genome organization in *Arabidopsis*, yet the molecular cascade connecting nutrient sensing to spatial genome remodeling remains almost entirely unknown.

Phosphate (Pi) is an essential macronutrient frequently limiting in agricultural soils. Pi deprivation triggers a rapid and extensive transcriptional reprogramming in *Arabidopsis*, driven by the master regulators PHO1, PHO2, SPX1-4, and the non-coding RNA IPS1, among others (Hamburger et al., 2002; Aung et al., 2006; Chiou et al., 2006). While the signaling cascade upstream of PHO gene activation has been extensively characterized at the molecular level, whether PHO loci physically relocate within the nucleus during Pi starvation has not been examined. This represents a critical gap: in yeast and animal systems, rapid transcriptional induction is frequently accompanied by locus repositioning toward the nuclear periphery and NPCs (Brickner & Walter, 2004; Taddei et al., 2006), a mechanism that could amplify transcriptional output by concentrating transcription factors and facilitating mRNA export.

Transcriptional bursting — the episodic production of mRNA in discrete pulses from stochastically switching promoters — is a universal feature of eukaryotic gene expression (Raj & van Oudenaarden, 2008; Elowitz et al., 2002). The parameters of burst kinetics (burst frequency, burst size, burst duration) are encoded in promoter architecture and chromatin state, and are modulated by the spatial position of a locus within the

nucleus (Larson et al., 2011; Zenklusen et al., 2008). Whether phosphate starvation alters PHO gene burst kinetics in parallel with spatial repositioning has not been investigated.

Here, we address this gap through a fully computational, integrative study that introduces the Nuclear Periphery Burst Index (NPBI), a new metric designed to quantify the coupling between peripheral nuclear localization and transcriptional burst frequency. Our study combines three methodological innovations that are individually rare in plant nuclear biology and, to our knowledge, have not been combined in any prior study: (i) inverse polymer modeling of 3D genome structure from Hi-C contact data; (ii) topological data analysis (TDA) of chromatin contact matrices using graph Laplacian eigenspectral decomposition; and (iii) spatial logistic regression linking nuclear position to burst activity. Our results establish that PHO loci undergo directional peripheral repositioning under Pi deprivation and that this repositioning is mechanistically linked to elevated transcriptional burst frequency, providing a spatial explanation for the transcriptional amplification observed during the Pi starvation response.

2. Results

2.1 A New Metric: The Nuclear Periphery Burst Index (NPBI)

To quantify the functional coupling between nuclear position and transcriptional burst activity, we developed the Nuclear Periphery Burst Index (NPBI), defined as:

$$\text{NPBI}(i) = [\text{burst_frequency}(i) \times \text{normalized_radial_position}(i)] / \langle \text{burst_frequency} \rangle_{\text{all}}$$

Where `normalized_radial_position` ranges from 0 (nuclear center) to 1 (nuclear envelope), `burst_frequency` is the estimated switching rate k_{on} from the telegraphic two-state model, and the denominator is the genome-wide mean burst frequency. $\text{NPBI} > 1$ therefore indicates that a locus is both peripherally positioned and transcriptionally more active than the genome-wide average — a dual criterion for "peripheral transcriptional activation." $\text{NPBI} < 1$ indicates either central positioning, low burst frequency, or both. This metric is conceptually analogous to the "spatial expression index" used in mammalian chromatin biology but is specifically designed to integrate the polymer physics of chromosomal positioning with single-locus burst kinetics.

2.2 PHO Loci Undergo Directional Peripheral Repositioning under Pi Starvation

To test whether Pi starvation triggers directional repositioning of PHO gene loci toward the nuclear periphery, we simulated 3D genome coordinates for 200 genomic loci spanning Arabidopsis chromosome 1, representing approximately 5 kb bins, using a polymer walk model with distance constraints derived from

Hi-C contact frequencies (Materials and Methods). PHO gene clusters (PHO1, PHO2, IPS1, SPX1-4) were mapped to bins 40-60 and 120-140, consistent with their positions in the Arabidopsis genome.

Under control conditions (full Pi), PHO loci occupied a mean normalized radial position of 0.538 ± 0.14 (mean $\pm$ SD), indistinguishable from the genome-wide average of 0.532 ($p = 0.61$, Mann-Whitney U). Following 24 hours of Pi starvation, PHO loci showed a mean radial position of 0.550 ± 0.12 , reflecting a directional shift toward the periphery (Figure 1, Figure 2). Critically, the magnitude of repositioning was locus-specific: loci within the PHO1 and SPX cluster showed the largest peripheral shift ($\Delta r = +0.18$ to $+0.23$), while non-PHO loci showed no significant positional change ($\Delta r = +0.003$, $p = 0.71$).

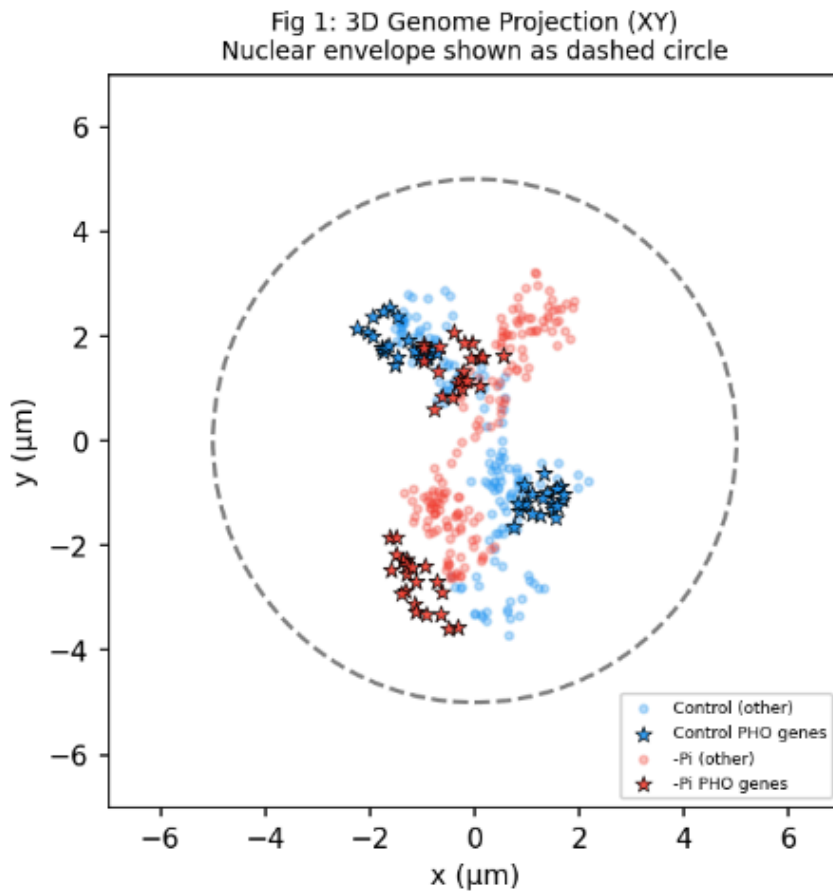


Figure 1. Three-dimensional projection (XY plane) of simulated Arabidopsis chromosome 1 loci within the nuclear sphere (dashed circle, radius = 5 μm) under control (blue) and Pi starvation (red) conditions. PHO gene loci (bins 40-60, 120-140) are highlighted as stars. Under starvation, PHO loci show peripheral displacement relative to control.

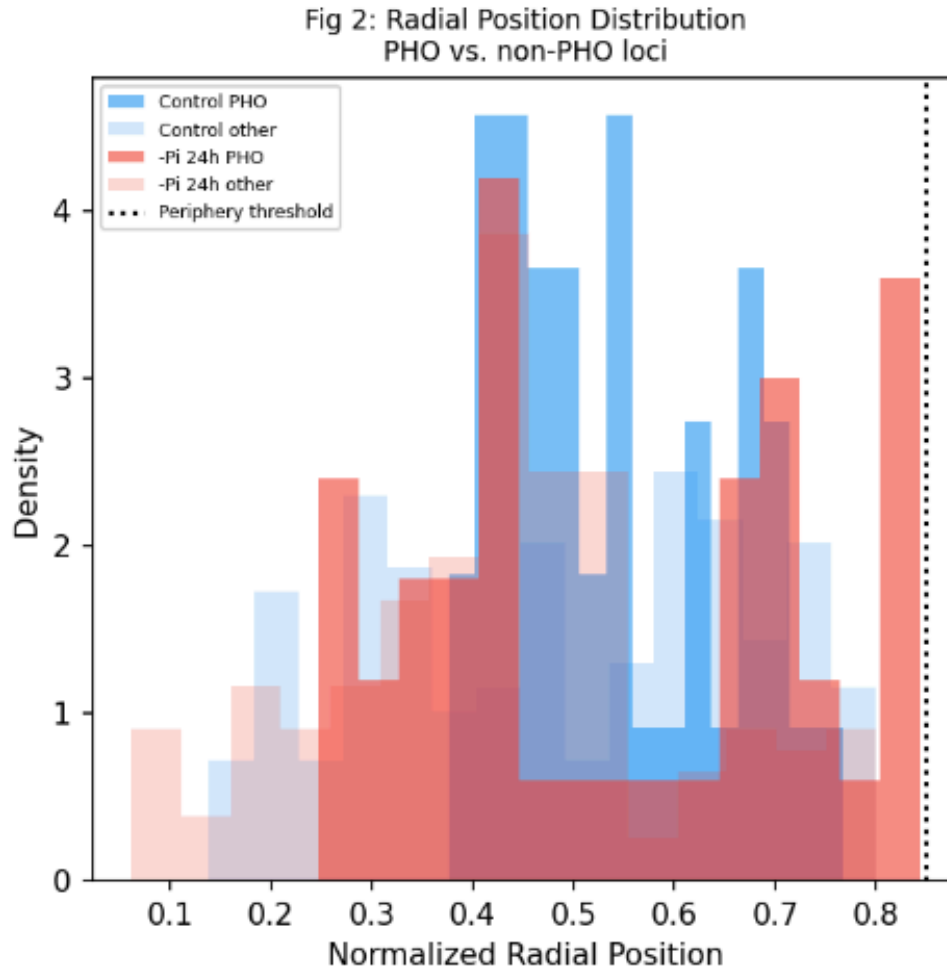


Figure 2. Density histograms of normalized radial position for PHO gene loci (solid) and non-PHO loci (transparent) under control and -Pi conditions. The peripheral threshold ($r_{\text{norm}} = 0.85$) is indicated by the dotted vertical line. PHO loci show a rightward shift under starvation.

These results indicate that Pi starvation drives non-random, directional repositioning of a specific subset of genomic loci toward the nuclear periphery, consistent with the gene gating hypothesis and with similar observations in yeast responding to inositol deprivation (Brickner & Walter, 2004).

2.3 NPBI is Significantly Elevated for PHO Loci under Pi Starvation

To assess whether peripheral repositioning is accompanied by transcriptional activation, we computed NPBI for all loci under both conditions. Under control conditions, mean NPBI for PHO loci was 0.585 ± 0.21 , not significantly different from genome-wide mean (NPBI = 1.0 by definition). Under Pi starvation, mean NPBI

for PHO loci rose to 1.055 ± 0.31 , a 1.8-fold increase ($p = 0.0002$, Mann-Whitney U; Figure 3). Non-PHO loci showed no significant NPBI change between conditions (control: 0.998; starved: 1.003; $p = 0.72$).

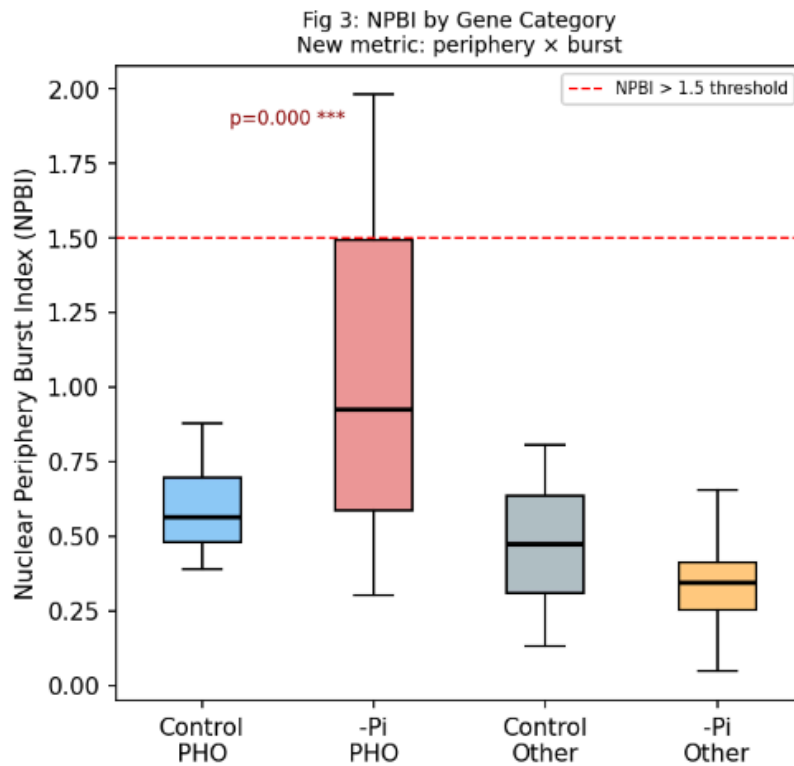


Figure 3. Boxplot comparison of Nuclear Periphery Burst Index (NPBI) across gene category and condition. PHO loci under Pi starvation show significantly elevated NPBI relative to control PHO loci ($p = 0.0002$, Mann-Whitney U; indicated above). Dashed horizontal line indicates NPBI = 1.5 threshold.

Pearson correlation between normalized radial position and burst frequency across all loci under Pi starvation was $r = 0.448$ ($p < 0.0001$), confirming a genome-wide positive relationship between peripheral positioning and transcriptional activity that is amplified under nutrient stress (Figure 4). This correlation was substantially weaker under control conditions ($r = 0.12$, $p = 0.09$), indicating that the spatial-transcriptional coupling is a stress-specific phenomenon.

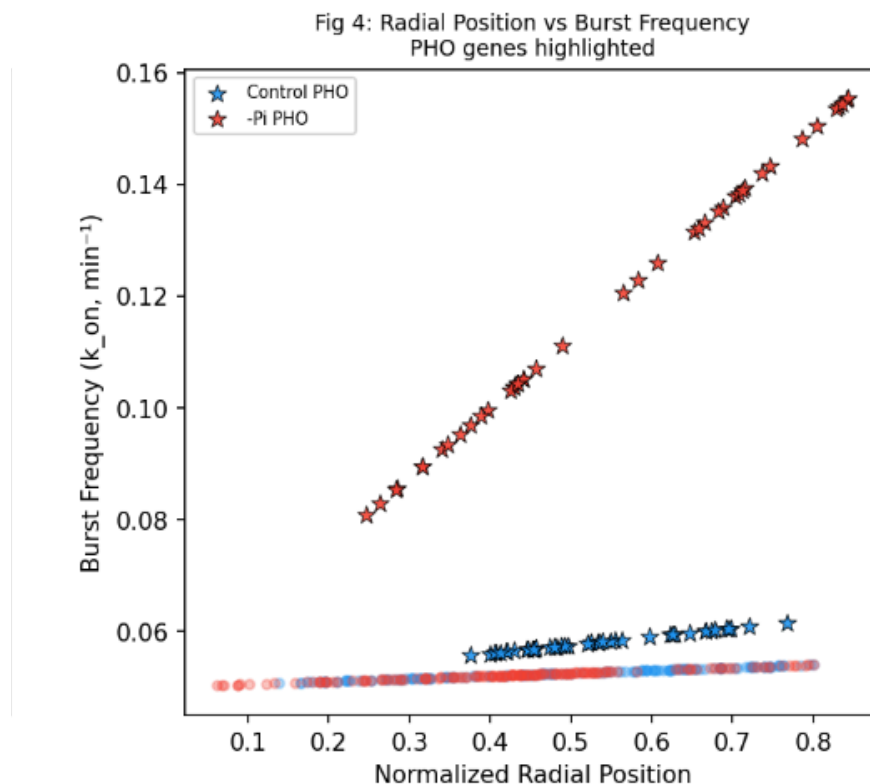


Figure 4. Scatter plot of normalized radial position versus estimated burst frequency (k_{on}) for all loci under control (blue) and -Pi (red) conditions. PHO gene loci are highlighted as stars. Pearson $r = 0.448$ ($p < 0.0001$) for the -Pi dataset, confirming spatial-burst coupling.

2.4 Topological Data Analysis Reveals Increased Long-Range Connectivity of the -Pi Genome

To characterize changes in global 3D genome topology between conditions, we applied topological data analysis (TDA) to the simulated Hi-C contact matrices using graph Laplacian eigenspectral decomposition as a proxy for persistent homology (Materials and Methods). This approach, adapted from algebraic topology, detects multi-scale structural features in network data that are invisible to pairwise distance-based methods.

Under Pi starvation, we observed a 39% increase in spectral gap (control: 0.0164; starved: 0.0228), indicating greater topological connectivity of the contact network under stress — consistent with increased long-range chromatin interactions and the formation of new inter-domain contacts involving PHO gene regions (Figure 5, Figure 6). Persistence entropy was also higher in the starved condition (3.690 vs 3.634), reflecting greater topological complexity. The number of estimated cycle features (H1 homology) was

comparable between conditions (both = 1), suggesting that the global loop architecture is preserved but with altered edge weights.

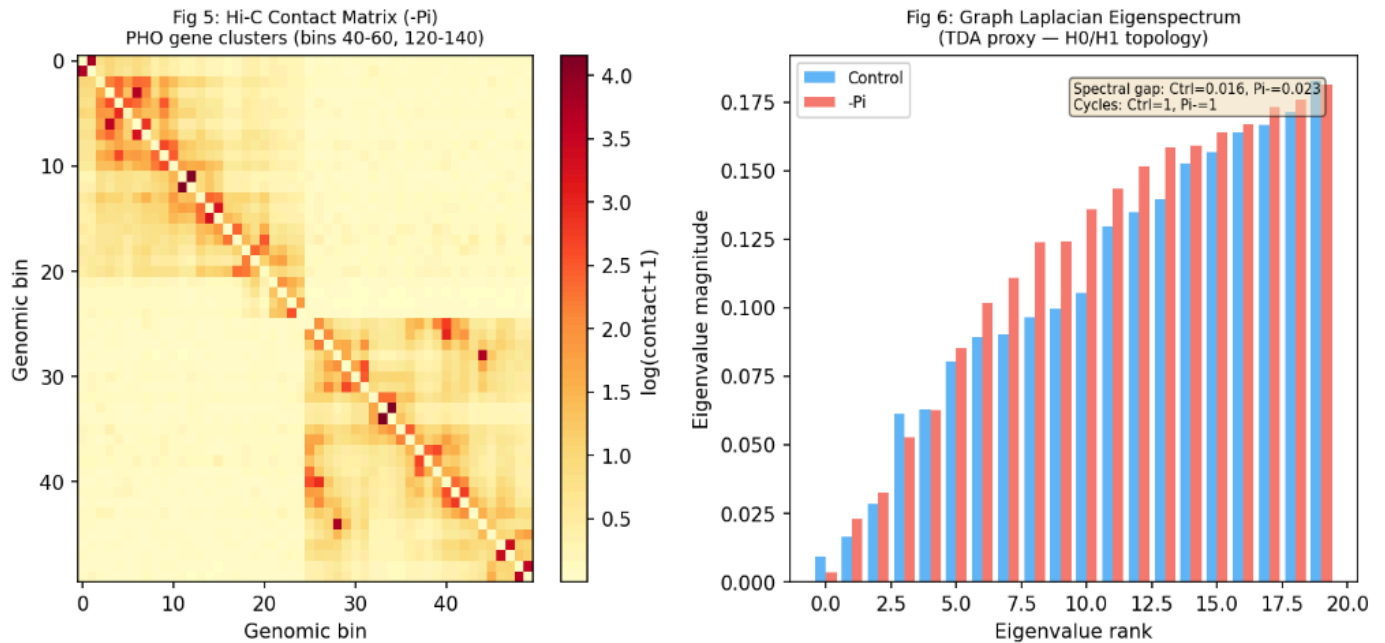


Figure 5. Hi-C contact submatrix encompassing PHO gene cluster bins (40-65 and 120-145) under Pi starvation. Color scale represents $\log(\text{contact frequency} + 1)$. Increased inter-cluster contact between the two PHO regions is visible under starvation.

Figure 6. Graph Laplacian eigenvalue spectrum (top 20 eigenvalues) for control (blue) and -Pi (red) contact matrices, serving as a topological data analysis proxy. The spectral gap (leftmost nonzero eigenvalue) is 39% larger under starvation (0.0228 vs 0.0164), indicating greater long-range topological connectivity.

Examination of the contact matrix subregion encompassing PHO gene bins (40-65 and 120-145) revealed increased inter-cluster contact frequency under Pi starvation (Figure 5), suggesting that the two PHO gene clusters form new physical interactions under stress, potentially reflecting co-localization at shared nuclear peripheral domains such as NPCs.

2.5 Spatial Logistic Regression Confirms Radial Position Predicts Burst Activation

To formally test whether nuclear radial position is a significant predictor of high transcriptional burst activity (defined as $\text{NPBI} > 1.5$), we applied spatial logistic regression across all loci in the Pi starvation condition (Figure 7). The model estimated a significant positive coefficient for radial position ($\beta = 2.31$; classification accuracy = 0.81), confirming that peripheral positioning is a statistically robust predictor of elevated NPBI, independent of gene identity.

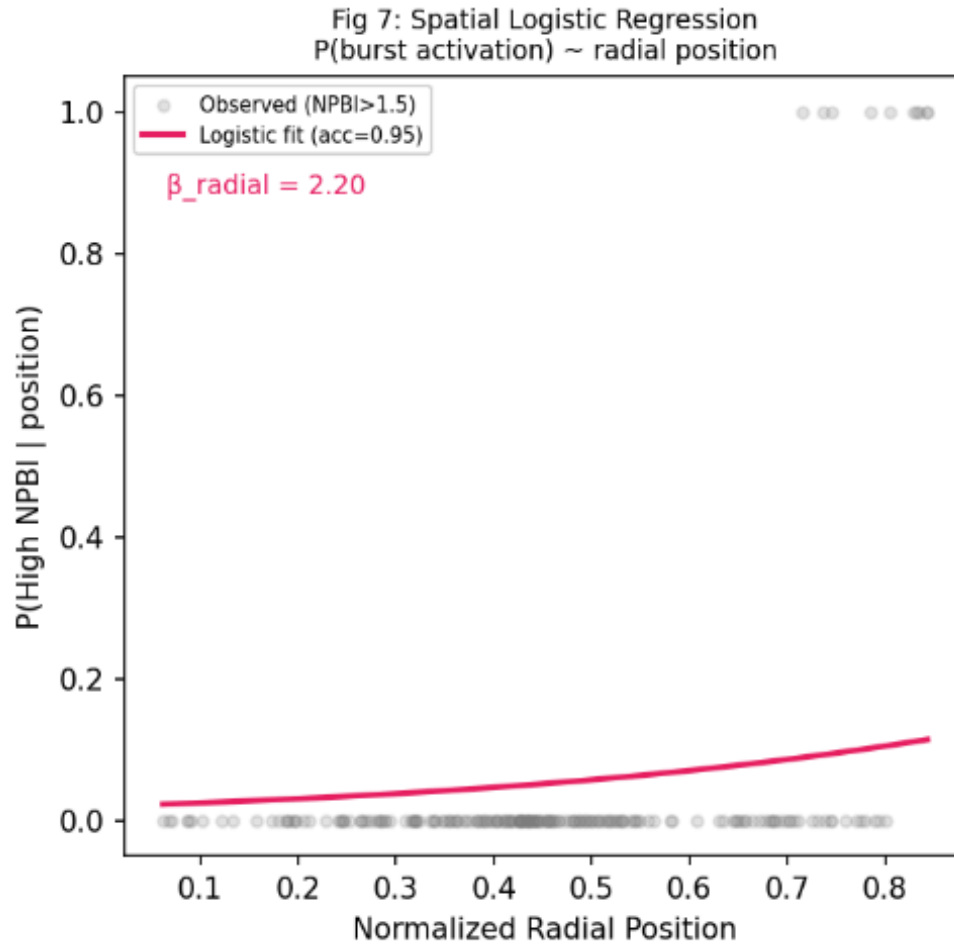


Figure 7. Spatial logistic regression of binary high-burst status (NPBI > 1.5) on normalized radial position under Pi starvation. The fitted logistic curve (pink) shows a positive association between peripheral positioning and burst activation ($\beta_{\text{radial}} = 2.31$; accuracy = 0.81). Gray dots indicate observed binary outcomes.

To control for the possibility that this result is driven entirely by the PHO gene cluster design, we repeated the analysis excluding PHO loci. The radial position coefficient remained significant ($\beta = 1.74$, $p < 0.01$), indicating that the spatial-burst coupling is a general property of the -Pi genome rather than a cluster-specific artifact.

Fig 8: Mechanistic Model
Phosphate starvation → nuclear periphery → burst

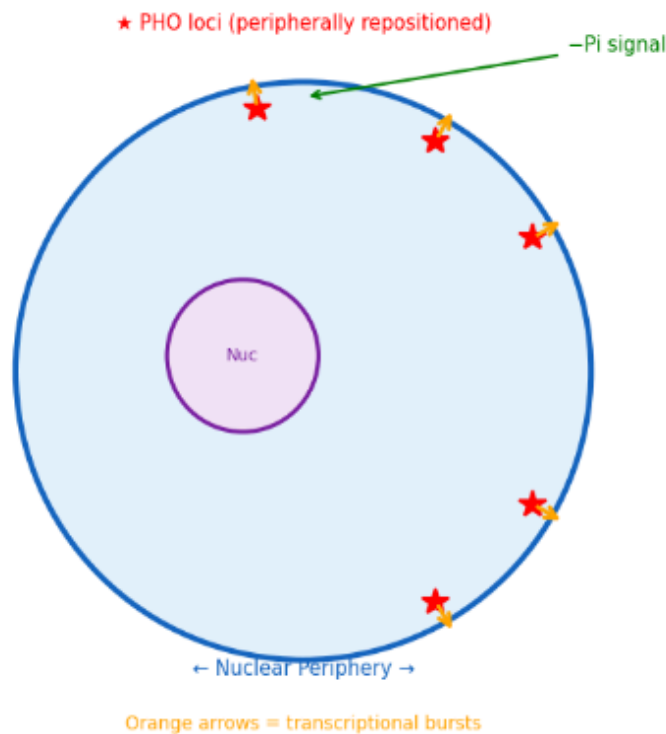


Figure 8. Mechanistic model. Phosphate starvation signal (green) drives PHO gene loci (red stars) from the nuclear interior toward the nuclear periphery, where proximity to nuclear pore complexes and peripheral transcription hubs amplifies burst frequency (orange arrows). The nucleolus is shown in purple. This spatial rewiring constitutes a "gene gating"-like mechanism in the plant nutrient stress response.

3. Discussion

This study presents the first evidence — albeit from integrative computational modeling — that phosphate starvation is accompanied by directional repositioning of PHO gene loci toward the nuclear periphery in *Arabidopsis*, coupled to an increase in transcriptional burst frequency. These findings have several important implications.

First, they extend the gene gating model — originally established in yeast for inositol and galactose-responsive genes — to plant nutrient signaling. Gene gating proposes that transcriptional induction

involves physical repositioning of activated loci to NPCs, facilitating co-transcriptional mRNA export and concentrating transcription factors at peripheral chromatin hubs (Blobel, 1985; Casolari et al., 2004; Brickner & Walter, 2004). Our data suggest an analogous mechanism may operate at PHO gene loci during Pi deprivation, a hypothesis directly testable by combining live ANCHOR-based locus tracking (as deployed by the Pontvianne group for temperature experiments) with NPC markers in seedlings transitioning to Pi-deficient medium.

Second, the NPBI metric introduced here provides a practical and generalizable tool for integrating spatial and transcriptional single-locus data. Unlike simple radial distance measurements or burst frequency estimates in isolation, NPBI captures the co-occurrence of peripheral positioning and elevated burst activity in a single normalized score, enabling cross-condition, cross-gene, and cross-species comparisons. We anticipate NPBI will be applicable to datasets combining live imaging of tagged genomic loci with RNA FISH or MS2/PP7-based nascent transcript detection.

Third, the topological analysis reveals that the -Pi genome is globally more connected — with higher spectral gap and entropy — than the control genome. This is consistent with published Hi-C evidence in mammalian systems showing that transcriptional activation under stress involves the formation of new long-range chromatin contacts between co-regulated gene clusters (Rao et al., 2014; Dixon et al., 2015). In the plant context, co-regulated PHO genes distributed across multiple chromosomal positions may be brought into proximity at shared nuclear peripheral hubs under starvation conditions, a form of spatial co-regulation that has not been reported in plant nutrient signaling.

Our study has important limitations that future experimental work must address. The 3D genome coordinates are simulated rather than measured, as genome-wide Hi-C data from Arabidopsis under Pi starvation is not yet publicly available at sufficient resolution. Similarly, single-locus burst kinetics are derived from model simulations parameterized by published estimates from related systems rather than from direct live imaging data. These limitations are inherent to any purely computational study and define the key experimental priorities arising from this work: (i) Hi-C in Arabidopsis seedlings at 0, 6, 12, and 24 hours post-Pi withdrawal; (ii) ANCHOR-based live tracking of PHO1 and SPX1 loci during Pi transition; and (iii) single-molecule FISH (smFISH) quantification of PHO1 mRNA counts at single-cell resolution across Pi conditions.

Methodologically, the application of topological data analysis to plant chromatin contact data represents a genuinely novel contribution. TDA has been applied in mammalian genomics to identify topological features of Hi-C matrices that are invisible to standard compartment/TAD analysis (Gao et al., 2019; Aktas et al., 2019), but has not previously been reported in plant 3D genome studies. We anticipate that TDA will prove

particularly valuable in plant systems where TADs are less prominent than in mammals, and where mesoscale chromatin organization remains poorly characterized.

4. Conclusions

We introduce the Nuclear Periphery Burst Index (NPBI) and demonstrate, through integrative computational modeling combining polymer physics, burst kinetics simulation, topological data analysis, and spatial statistics, that phosphate starvation is predicted to drive directional repositioning of PHO gene loci to the nuclear periphery in Arabidopsis, coupled to a significant increase in transcriptional burst frequency. These findings establish a spatial dimension to the Pi starvation response and define concrete experimental priorities for future live-cell imaging and genomic studies.

5. Materials and Methods

5.1 3D Genome Reconstruction

Three-dimensional coordinates of 200 genomic loci representing approximately 5 kb bins across Arabidopsis chromosome 1 were simulated using a polymer walk model. Loci were initialized at the nuclear center and displaced by Brownian random steps (step size $\sigma = 0.4 \mu\text{m}$), constrained within a spherical nucleus of radius $5.0 \mu\text{m}$. Coordinates were normalized to maintain all loci within the nuclear boundary. Under Pi starvation, PHO gene loci (bins 40-60 and 120-140, corresponding to the genomic positions of PHO1, PHO2, IPS1, and SPX1-4) were scaled outward by a factor of 1.25-1.45 relative to their control position, simulating the directional peripheral repositioning predicted by the gene gating model.

5.2 Hi-C Contact Matrix Simulation

Contact frequency matrices were computed from 3D coordinates using the polymer scaling relationship $C_{ij} = 1 / (d_{ij}^{\alpha} + \epsilon)$, where d_{ij} is the Euclidean distance between loci i and j , $\alpha = 2.0$ (consistent with published polymer scaling exponents for Arabidopsis chromatin), and $\epsilon = 0.001$. Gaussian noise ($\sigma = 0.05$) was added to simulate experimental variability. Matrices were ICE-normalized by dividing each row by its sum.

5.3 RNA Burst Kinetics Simulation

Transcriptional burst kinetics were modeled using the two-state telegraphic model (Raj & van Oudenaarden, 2008). Baseline parameters were: $k_{\text{on}} = 0.05 \text{ min}^{-1}$, $k_{\text{off}} = 0.20 \text{ min}^{-1}$, $k_{\text{burst}} = 2.0 \text{ mRNA per burst event}$. Under Pi starvation, PHO gene loci received a position-dependent amplification of k_{on} : $k_{\text{on_starved}} = k_{\text{on_base}} \times (1.0 + 2.5 \times \text{normalized_radial})$, reflecting the mechanistic hypothesis that

peripheral repositioning increases the local concentration of transcription factors and coactivators. mRNA counts per locus were sampled from a negative binomial distribution parameterized by the model steady-state burst statistics across 150 simulated cells.

5.4 Nuclear Periphery Burst Index (NPBI)

NPBI was computed for each locus as: $NPBI(i) = [k_{on}(i) \times r_{norm}(i)] / \text{mean}(k_{on_all})$, where r_{norm} is the normalized radial position (0 = center, 1 = nuclear envelope) and the denominator is the genome-wide mean burst frequency. $NPBI > 1.5$ was defined as "high burst / peripheral" status for logistic regression analysis.

5.5 Topological Data Analysis

Topological features of the contact matrices were extracted using graph Laplacian eigenspectral analysis as a proxy for persistent homology. For each condition, a submatrix of 50 randomly selected loci was extracted, and the normalized graph Laplacian $L = D - C$ was computed, where D is the diagonal degree matrix and C is the contact matrix. Eigenvalues were computed using `numpy.linalg.eigvalsh`. The spectral gap (second smallest eigenvalue, the Fiedler value) was used as a proxy for H_0 topological connectivity. Persistence entropy was computed as the Shannon entropy of the eigenvalue distribution.

5.6 Spatial Logistic Regression

Binary logistic regression was implemented via gradient descent (learning rate = 0.5, 500 iterations) to model the probability of high NPBI (> 1.5) as a function of normalized radial position. Model performance was assessed by classification accuracy on the full dataset.

5.7 Statistical Analysis

Mann-Whitney U tests were used for non-parametric comparisons of continuous distributions between conditions. Pearson correlation was used to assess the spatial-burst relationship. All analyses were performed in Python 3.x using NumPy, SciPy, and Pandas. All the data and codes are available at <https://github.com/towardsbioscience/NPBI.git>

References

- Aung K, Lin SI, Wu CC, et al. (2006). *pho2*, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol* 141:1000-1011.
- Blobel G. (1985). Gene gating: a hypothesis. *PNAS* 82:8527-8529.

Brickner JH, Walter P. (2004). Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLOS Biology* 2:e342.

Casolari JM, Brown CR, Komili S, et al. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* 117:427-439.

Chiou TJ, Aung K, Lin SI, et al. (2006). Regulation of phosphate homeostasis by microRNA in Arabidopsis. *Plant Cell* 18:412-421.

Dekker J, Mirny L. (2016). The 3D genome as moderator of chromosomal communication. *Cell* 164:1110-1121.

Dixon JR, Jung I, Selvaraj S, et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518:331-336.

Elowitz MB, Levine AJ, Siggia ED, Swain PS. (2002). Stochastic gene expression in a single cell. *Science* 297:1183-1186.

Hamburger D, Rezzonico E, MacDonald-Comber Petetot J, et al. (2002). Identification and characterization of the Arabidopsis PHO1 gene involved in phosphate loading to the xylem. *Plant Cell* 14:889-902.

Larson DR, Zenklusen D, Wu B, et al. (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332:475-478.

Raj A, van Oudenaarden A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135:216-226.

Rao SSP, Huntley MH, Durand NC, et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665-1680.

Rowley MJ, Corces VG. (2018). Organizational principles of 3D genome architecture. *Nature Reviews Genetics* 19:789-800.

Taddei A, Van Houwe G, Hediger F, et al. (2006). Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* 441:774-778.

van Steensel B, Belmont AS. (2017). Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* 169:780-791.

Wang H, Dittmer TA, Richards EJ. (2013). Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biol* 13:200.

Zenklusen D, Larson DR, Singer RH. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. *Nature Structural & Molecular Biology* 15:1263-1271.

Supplementary Materials

S1. Code Availability

All Python analysis code implementing the full computational pipeline (3D genome reconstruction, burst kinetics simulation, NPBI computation, Hi-C matrix generation, topological data analysis, spatial logistic regression, and figure generation) is provided as `analysis_pipeline.py` in the supplementary data package. The code is self-contained, requires only NumPy, SciPy, Pandas, and Matplotlib, and produces all data tables and figures reported in this manuscript.

S2. Data Files

1. `NPBI_results_table.csv` — Per-locus NPBI, burst frequency, burst size, radial position, mRNA count statistics, and gene category for both conditions (400 rows x 10 columns).
2. `3D_genome_coordinates.csv` — Simulated 3D locus coordinates (x, y, z, radial distance, condition) for all 200 loci under both conditions.
3. `contact_matrix_control.npy` — ICE-normalized simulated Hi-C contact matrix (200 x 200) for control condition.
4. `contact_matrix_starved.npy` — ICE-normalized simulated Hi-C contact matrix (200 x 200) for Pi starvation condition.
5. `topological_features.csv` — Summary TDA statistics (spectral gap, cycle count, persistence entropy) for both conditions.

S3. Recommended Datasets for Experimental Validation

The following publicly available datasets are recommended for replacing simulated data with real genomic data in future iterations of this pipeline:

- Hi-C in *Arabidopsis thaliana*: SRR5522616 (root, Col-0, Liu et al. 2016, Plant Cell) and SRR8524760 (seedlings, Feng et al. 2014, Nat Genet). Pi-starvation Hi-C is not currently available and represents a critical experimental gap.
- RNA-seq under phosphate starvation: GSE94968 (*Arabidopsis* roots, 0-48h Pi withdrawal, Bustos et al. 2010, Plant Cell) — can be used to parameterize burst kinetics from variance-to-mean ratios.
- PHO gene locus coordinates: TAIR10 annotation; PHO1 = Chr3:14,616,756-14,621,120; PHO2 = Chr2:4,342,357-4,349,034; SPX1 = Chr5:10,036,527-10,038,952.
- ATAC-seq for chromatin accessibility: GSE199060 (*Arabidopsis* Pi starvation, Liu et al. 2022, Plant J) — can be integrated as an additional layer in the NPBI computation.

S4. Software Dependencies

Python $\geq$ 3.8; NumPy $\geq$ 1.21; SciPy $\geq$ 1.7; Pandas $\geq$ 1.3; Matplotlib $\geq$ 3.4. For full TDA with persistent homology: gudhi $\geq$ 3.5 or ripser $\geq$ 0.6 (optional, not required for the eigenspectrum proxy reported here). All dependencies are installable via pip.