Two kinds of time in mammalian promoters: a refractory class that detects change faster per burst cycle

Matthew D. Schwartz*[1] and Martin Hemberg[2]

[1] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*
[2] *The Gene Lay Institute of Immunology and Inflammation, Brigham and Women's Hospital, Massachusetts General Hospital, and Harvard Medical School, Boston, MA, USA*

Correspondence: Matthew D. Schwartz (schwartz@g.harvard.edu); Martin Hemberg (mhemberg@bwh.harvard.edu)

This manuscript is preliminary. The first draft was written by Claude (Anthropic), and the text and computations are still under review by the authors.

September 28, 2026

Abstract

Transcription in mammalian cells occurs in bursts, and the standard promoter model is a two-state switch with memoryless silent periods. Fitting exact likelihoods for that model and a multi-step alternative to time-resolved single-cell RNA counts from primary mouse fibroblasts, we find that a subset of genes, enriched five-fold for secreted and matrix proteins, builds its silences from several sequential steps. This refractory class is CpG-poor, and no transcription-factor motif is enriched at its promoters beyond base composition. Its promoters are methylated yet carry the active marks H3K4me3 and H3K27ac at reduced levels, with more H3K27me3 than memoryless promoters. We develop a mathematical model suggesting that the multi-step promoter architecture enables detection of a change in its input several-fold sooner, measured in its own burst cycles, than a memoryless switch.

Keywords: transcriptional bursting; promoter kinetics; refractory period; single-cell RNA sequencing; metabolic labeling; DNA methylation; chromatin; change detection; telegraph model

Introduction

In both prokaryotic and eukaryotic cells genes are not transcribed continuously[1–7]. A promoter spends most of its time silent and produces its transcripts in bursts, and the distribution of RNA counts across single cells carries the timing of that switch[3,5]. The standard model, often referred to as the telegraph model, treats the promoter as a two-state switch[8]: it turns on at a constant rate from a single silent state and off at a constant rate from a single active state, so its silent periods are exponentially distributed and memoryless. Several studies have used traditional single cell RNAseq (scRNAseq) data to fit this model[6,9–14], even though it has been shown that it cannot constrain the shape of the silent period. However, experiments in which newly synthesized transcripts are metabolically labeled with 4-thiouridine and read out by nucleotide conversion contain the required information, and several protocols have

*Work done with Anthropic, but results and views are not endorsed by Anthropic.

been developed, including scSLAM-seq[15], NASC-seq[16], sci-fate[17], scNT-seq[18] and NASC-seq2[19]. This telegraph model underlies most studies inferring burst kinetics from scRNAseq, including the genome-wide kinetic maps of mouse embryonic fibroblasts[6] and embryonic stem cells[12], as well as recent work relating burst kinetics to chromatin architecture, BRD4, and Mediator[20].

Whether one silent state is enough has been asked since the first single-cell time series: Golding and colleagues, following tagged transcripts in living *E. coli*, found silent periods that were exponentially distributed, as a single first-order step predicts[1], and the first eukaryotic pulses, in *Dictyostelium*, fit exponentials as well[2]. Studying single mouse fibroblasts using luminescent reporters Suter et al found silent periods too regular for a single first-order step, leading to the conclusion that a gene must pass through a refractory period before it can burst again[5]. Harper and colleagues independently made a similar observation for prolactin transcription[21], and Zoller et al fitted several rate-limiting steps to the silent intervals of a panel of mammalian genes[22]. Rodriguez and colleagues, following the human TFF1 gene in living cells, found the opposite deviation: fast cycling between on and off interrupted by a rare, long repressive state, so that silences are more variable than exponential and bursts arrive in clusters[23]. Similarly, a study of macrophages suggested that the secreted cytokine IL1β needs an extra, chromatin-permissive step while TNFα can be modeled using a single step[24]. Silences that mix a fast and a slow return have since been seen in other live-cell records: the non-permissive periods of an HIV-1 reporter are bi-exponential, with the two-timescale mixture fitting where a single exponential departs[25], and the interburst intervals of a developmental enhancer's target need three exponential components and a promoter that switches between a low- and a high-frequency regime[26].

While models based on time-series of either mRNA or proteins typically result in non-Markovian models, genome-scale inference from snapshots of single-cell counts has generally kept the two-state model[6,20]. There are structural reasons behind this: a stationary count distribution constrains the silent-period distribution mainly through its mean, and thus a snapshot is poorly suited for characterizing promoter memory[27,28]. Stinchcombe et al showed that in some parameter regimes exponential and non-exponential dwell times are indistinguishable from the count histogram alone[29]. They judged that "attempts to use transient temporal data following the induction of transcription would likely be unsuccessful due to the difficulty in obtaining the sufficient data during a single transient window per cell", and instead considered the likelihood of temporal sequences of counts under alternative dwell-time densities[29]. In a 2020 review of transcriptional bursting, Rodriguez and Larson wrote that the two-state telegraph model "is the de facto description for mRNA expression" although live-cell studies indicate it "does not quantitatively describe bursting dynamics for all genes"[30]. Inferring mechanistic models from single-cell counts through the two-state model goes back to the early single-cell RNA-seq fits, which asked whether a change in expression came from burst size or burst frequency[31]. Those fits evaluated the Poisson-Beta likelihood by Monte Carlo because the confluent hypergeometric function could not be computed reliably[31]. Time-resolved single-cell profiling carries more information than a snapshot, but these assays are low-throughput and each experiment characterizes at most a handful of genes.

In NASC-seq2 and similar protocols, newly synthesized RNA in each cell is labeled with 4-thiouridine for a defined window before sequencing and read out through T-to-C conversions, so that the new-RNA count of a gene reports its transcription during the window[19]. This experimental setup turns the inference problem from a stationary one into a transient one, and the transient distribution of new-RNA counts, unlike the stationary distribution, depends on the shape of the silent-period distribution. For the two-state model under a labeling window, Ramsköld and colleagues evaluated this distribution exactly, and fitted each gene to these data[19], although they did not consider any alternatives to the telegraph model. We derived exact likelihoods for the two-state model and for its multi-step generalization under a labeling window, with allele-resolved counts handled by exact convolution. The likelihoods were fitted to every gene on each of the two chromosomes separately, and the p-value of each gene's fit was computed exactly rather than by simulation (Methods). We then asked three questions in turn: which promoters need a third state; what the promoters that need it have in common, in function, in chromatin, and in

the machinery that serves them, and what the multi-step architecture does for the cell. The class that needs a third state includes 186 genes, or 15 percent of the 1,246 genes where we had enough data to conclusively distinguish between the two models (1.1 percent of the 16,858 genes where the fits for both models converged), is enriched 4.8-fold for secreted and extracellular-matrix proteins, sits on CpG-poor methylated promoters, and keeps silences of days built from several sequential steps, with the length of the silence tracking promoter methylation. Secreted and matrix proteins are the gene category in which days-long repressive periods[23] and stable, low-level allelic silencing[32–34] have been described. We use publicly available chromatin data to exclude a polymerase pause as the extra state, consistent with off-states timed by chromatin rather than by polymerase. To investigate the benefits of a refractory promoter we develop a mathematical model which suggests that it enables several-fold faster detection of a change in the promoter's input, measured in the promoter's own burst cycles. We close with a test of generality on the X chromosome and outline several experiments to test the predictions of our model.

Results

Exact fits to time-resolved counts divide fibroblast promoters into two opposite kinds of temporal structure

We consider RNA counts corresponding to newly synthesized transcripts from 8,916 primary mouse fibroblasts after a two-hour 4-thiouridine window[19]. Importantly, the cells are F1 hybrids, so counts are resolved by allele. We use this data to compare the telegraph model with a refractory model involving multiple states for the inactive promoter (Figure 1A), In the two-state model the promoter leaves its silent state in a single first-order step, so its silent periods follow an exponential distribution. In the multi-step model it returns from silence through a chain of k sequential first-order steps, so its silent periods have lower variance and a refractory period after each burst; the active state and the synthesis and degradation of RNA are the same in both models. For our experimental design with a labeling window, the two models predict the same mean number of new-mRNA counts, but as the distributions differ the exact transient likelihood measures will also differ. We fitted both models by maximum likelihood to the 23,246 detected genes whose count range fit within the likelihood solver (Methods), with degradation rates obtained from another study[6], and obtained converged fits of both models for 16,858 genes, which we call evaluable. Of the 6,388 genes where both alleles did not converge, 4,028 of them neither allele converged, and these are low-count genes (median 53 new molecules across all cells against 928 for the evaluable genes). For the remaining 2,360 one of the two optimizers stopped short of the convergence criterion (Methods). For each evaluable gene we used an exact test, described in the companion paper[28], to compute the p-value of the likelihood-ratio statistic exactly from the fitted two-state model (16,844 genes) or bounded it with a Chernoff bound where the exact computation was out of reach (14 genes).

Using the telegraph-model as a null hypothesis, this procedure rejects it for 3,737 genes at a false-discovery rate of 5 percent[35] on the exact test alone (Figure 1B). To determine how many of these genes can be confidently assigned to the multi-step model, we carried out three additional tests. First, we tried to reproduce the observed statistic without a third state: for each gene we simulated counts from its own fitted two-state model at its own sequencing depth, with per-cell capture dispersion, and refitted both models to the simulated counts. Second, we repeated this with a model which accounts for the copy number differences arising due to the cell cycle. Third, we restricted the fits to cells in the G1 phase of the cell cycle to ensure that replicated gene copies cannot inflate the statistics (Methods). This left us with 186 genes whose likelihood ratio statistic exceeds three times each of the two refitted models, and which are also assigned to the multi-step model when only the G1 cells are used. In the other direction, 1,060 genes favor the two-state model by more than the same margin (statistic below -4.3). The refractory genes therefore constitute 15 percent of the 1,246 genes assigned to either class and 1.1 percent of the 16,858 evaluable genes. The remaining 15,612 evaluable genes (92.6 percent)

are undetermined at this depth of data rather than assigned to either class (Figure 1C). The 1,246 high confidence genes will be used for the remainder of the paper. To evaluate the robustness of this result, we considered several alternative filtering criteria, using either more relaxed or more stringent thresholds (Figure S1). Requiring a 2x rather than 3x margin over the two refitted models enlarges the refractory class from 186 to 305 genes. Requiring any two of the three controls gives 374 genes, and both relaxations together give 574, while a stricter 4x margin keeps 127 (68 percent) (Methods; Table S5) (Figure 1E). Across all 8,916 cells, refractory genes are expressed at a median of 0.22 new transcripts per cell (interquartile range 0.10 to 0.49), 6.2-fold below the 731 telegraph genes (median 1.36; Mann-Whitney on the log mean, $p = 3 \times 10^{-69}$) and 2.9-fold above the 15,612 undetermined genes (median 0.075; $p = 2 \times 10^{-14}$). A refractory gene is detected in a median of 14 percent of cells, against 63 percent for telegraph genes. Every classed gene has a Fano factor greater than 1, and at matched mean refractory genes carry 1.5-fold more cell-to-cell variance than the genome-wide trend (Figure S2). The gap in expression is partly a selection effect, since a confident class call needs enough counts to reject the alternative model, and consequently we can conclude that the refractory class is not associated with higher counts.

The refractory model shows that some genes have longer off-state than allowed for by the telegraph model, and we also investigated if there are some genes that show evidence of off-states clustered into two groups as suggested by experimental data[26]. We revisited the NASC-seq2 data to compare a two-state model and fitted a clustered promoter, whose silent period is a two-timescale mixture with a fixed shape (squared coefficient of variation 2, Figure 1A). Because the clustered promoter is a refinement of the memoryless one, keeping its mean silent period but drawing each silence from two timescales instead of one, we call the clustered class only within the 1,060 genes of the two-state class. Promoters were evaluated using the same rule as before: the likelihood-ratio statistic must exceed 8.6 and exceed three times the larger of 4.3 and each of two competing models (a two-state gene refit under the same misspecification, and a two-state gene with the copy-number variation of the cell cycle), and the call must be better than a null model fitted only using G1 cells. Of the 1,060 two-state genes, 329 (31 percent) pass the bar, and we call these clustered. We used the same robustness test as for the refractory class (Figure S1B), and we found that requiring a 2x rather than 3x margin over the two refitted models enlarges the clustered class from 329 to 557 genes, while requiring any two of the three controls gives 453, and both relaxations together give 684. A stricter 4x margin keeps 235 (71 percent), and using the likelihood-ratio statistic alone, without any control, 744 of the 1,060 two-state genes were classed as clustered (Table S5). In summary, of the 1,246 high confidence genes, 186 were classed as refractory with a multi-step silent period, 329 clustered genes are better described by an exponential mixture, and 731 telegraph genes are memoryless (Methods).

With only a two hour labeling window it is challenging to infer the properties of the off-state distributions and to determine the number of steps for the multi-step model. To further investigate this aspect we took twenty genes with small count ranges, twelve from the refractory class and eight from the two-state class, spaced evenly along the ranked likelihood-ratio statistic within each class. We fitted each with a wider set of promoter models: chains of depth 3 and 5, the clustered (two-timescale) silent period, and a continuous profile over chain depth for the refractory genes (Figure 1D) or over the dispersion of the silent period for the two-state genes. For the twelve refractory genes the likelihood rises from the two-state model through depths 2, 3 and 5 for all twelve (twice the log-likelihood gain at depth 5: 17.6 to 88.0, median 30.0). The profile maximum sits at the largest depth of the primary grid, 8, for all twelve. When the grid was extended to depths 16 and 32 for eleven of them, the profile was flat beyond depth 8 for nine, flat beyond 16 for one, and still rising at 32 for one, so the fitted depths are lower bounds. For the eight two-state genes the deviation runs the other way: every chain fits worse than the two-state model, and the deficit grows with depth for all eight (blue lines in Figure 1D; twice the log-likelihood relative to the two-state model -0.04 to -17.2 at depth 2 and -0.2 to -34.9 at depth 5, median -7.6; for two of the eight the decrease is smaller than one unit). Twelve more clustered genes, chosen to

span the class from its weakest to its strongest call, behave the same way: every chain fits worse than the two-state model for all twelve, and their curves fall faster than the telegraph curves at the median (depth-5 deficit 12.1 against 7.4) without separating from them (one-sided Mann-Whitney $p = 0.07$; green lines in Figure 1D). Interestingly, a silent period with two timescales, fast cycling interrupted by occasional long silences, fits better for all eight (twice the log-likelihood gain 3.4 to 37.5, median 7.5). The profile over the dispersion of the silent period peaks above the exponential value for seven of the eight (squared coefficient of variation 1.5 to 1.8) and excludes the exponential at 95 percent confidence for five, which is the signature of burst clustering. For any renewal process the shape of the interval distribution identifies the kind of kinetic scheme with a single Poisson step giving a squared coefficient of variation of 1, a multi-step process less than 1, and a multi-channel process more than 1 [22,36,37]. Thus, there are three types of fibroblast promoters and they are either more regular than a memoryless switch, less regular than one, or consistent with the exponential promoter of the standard model.

The refractory class is enriched for secreted and matrix proteins and is silent for days

Next, we investigated the features of the 186 refractory promoters to gain insight into the mechanisms underlying this behavior. Gene Ontology analysis revealed that they are enriched 4.8-fold for the term extracellular region (62 of 164 annotated refractory genes, $p = 8.7 \times 10^{-24}$ after Benjamini-Hochberg correction) and 7.9-fold for extracellular matrix (25 of 164, $p = 7.5 \times 10^{-13}$), with collagens, matrix modifiers, and secreted signaling proteins prominent in the list (Figure 2A). Interestingly, Rodriguez and colleagues had already singled out secreted and signal-peptide genes as the category with long stochastic repressive periods, inferred from their extreme expression heterogeneity in tissue [23]. Our analysis extends that association to also include regular, multi-step silences. In primary fibroblasts of this same cross the small set of genes whose alleles are stably silent in clones is likewise low-expressed and enriched for secreted and extracellular proteins [32], as it is in the genome-wide surveys of stably silenced genes [33,34]. Three of the 186 refractory genes (Cd59a, Npnt, Psmb8) are on that clonal list for this cell type, against four of the 1,060 two-state genes. Moreover, refractory promoters are CpG-poor, and compared to expression-matched two-state promoters, refractory promoters carry 31 percentage points more CpG methylation (class means 43 against 12 percent over the 104 and 84 promoters with measured CpGs; Mann-Whitney $p = 1 \times 10^{-14}$). They overlap CpG islands 42 percentage points less often (39 against 81 percent; Fisher's exact test $p = 1 \times 10^{-16}$), whereas the two-state class is dominated by CpG-island promoters.

The 329 clustered genes carry a different functional signature. The same Gene Ontology analysis finds enrichment for regulators of cell differentiation (Figure 3E; 2.9-fold, $p = 2.0 \times 10^{-14}$), of cell population proliferation (3.0-fold; $p = 2.0 \times 10^{-14}$) and of intracellular signal transduction (2.4-fold; $p = 1.2 \times 10^{-10}$). It is also enriched for DNA-binding transcription factor activity (2.2-fold, $p = 1.3 \times 10^{-3}$) and growth factor binding (6.3-fold, $p = 9.8 \times 10^{-4}$). Inducible and immediate-early transcription factors (Jun, Egr1, Egr2, Klf2, Klf4, Cebpb, Stat3, Nfkb1, Bhlhe40, Nfil3, Sox4, Sox9) and growth-factor receptors and ligands (Pdgfrb, Tgfbr3, Notch1, Vegfa, Pdgfa, Kitl, Osmr, Il4ra) are prominent in the list. This signature belongs to the clustered subset rather than to the two-state class as a whole: the 731 telegraph genes are enriched instead for nuclear proteins (nucleus 1.5-fold, $p = 3.6 \times 10^{-19}$) and for regulation of primary metabolic process (1.5-fold; $p = 1.5 \times 10^{-11}$). The clustered genes are enriched 2.2-fold over telegraph genes for regulation of cell differentiation (76 of 299 against 72 of 620; Fisher's exact test, $p = 2.7 \times 10^{-7}$). In terms of sequence features the clustered genes are indistinguishable from telegraph genes: against expression-matched telegraph promoters (329 pairs) they differ by 1.5 percentage points in CpG methylation ($p = 0.72$) and there is no difference in CpG-island overlap (78 against 77 percent; $p = 0.85$) or length (median 40.0 versus 34.7 kb, P = 0.22).

The most conspicuous difference between the two-state class and the refractory class is the duration of the inactive periods (Figure 2B–D). The rate of leaving the silent state is 24-fold lower for the refractory

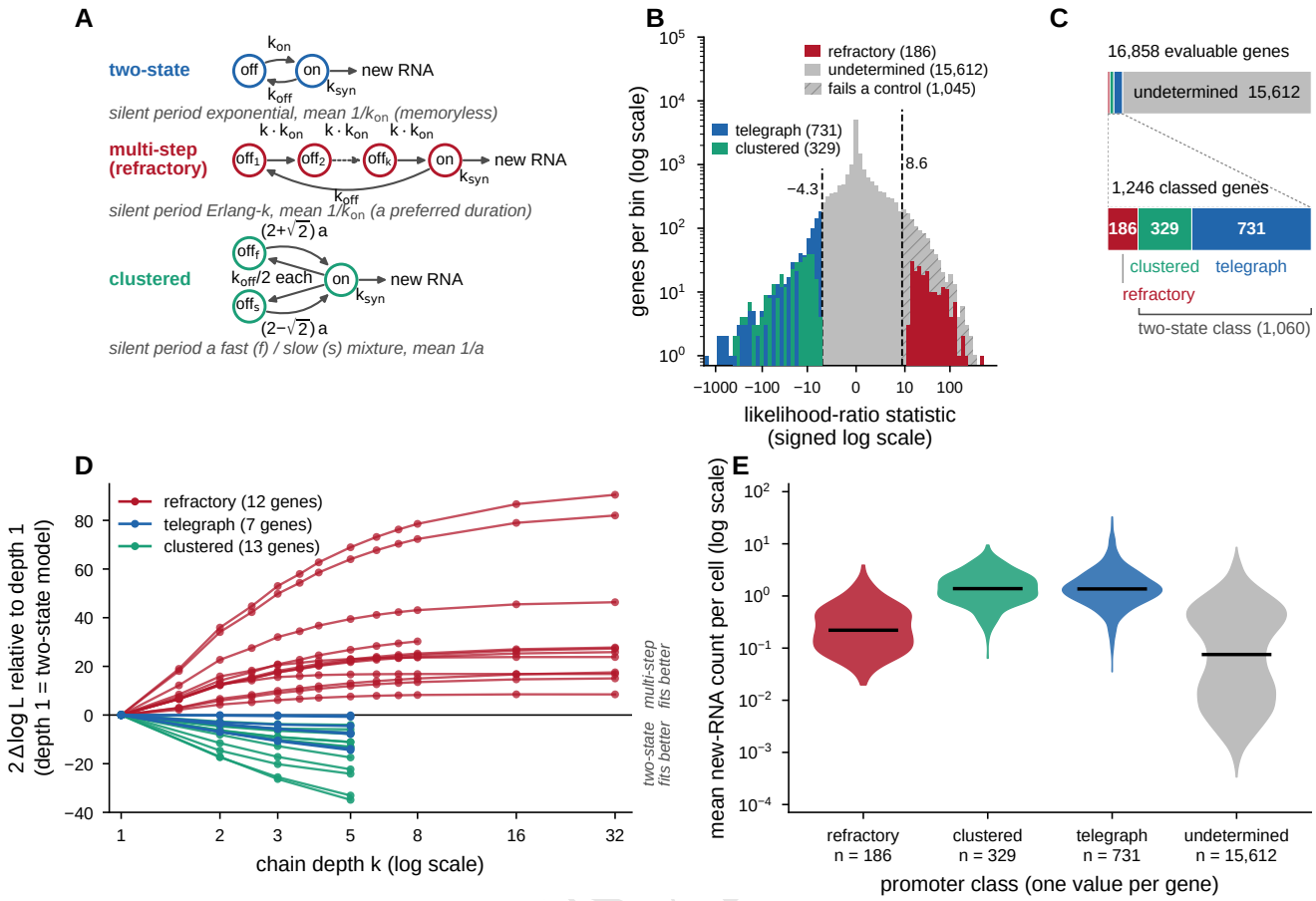

Figure 1: Exact fits to time-resolved counts divide mouse embryonic fibroblast promoters into two kinds of temporal structure. (A) The three promoter models. In the two-state model the promoter leaves silence in one first-order step; in the multi-step model it returns through k sequential steps, giving silent periods a preferred duration. In the clustered model it enters a fast or a slow silent state with equal probability after each burst, so silences are a two-timescale mixture with mean $1/a$ and bursts arrive in clusters. Synthesis during the labeling window produces the new-RNA count that is measured. (B) Likelihood-ratio statistic (multi-step against two-state, signed logarithmic axis) for the 16,858 evaluable genes, colored by class (refractory red, telegraph blue, clustered green, undetermined grey; histograms overlaid, not stacked; the count axis is logarithmic because the classes are a few percent of the peak bin). The 186 refractory genes exceed the threshold of 8.6 and pass three control constructions; the hatched grey bars above 8.6 are the 1,045 genes that exceed the threshold but fail a control; the 1,060 two-state genes favor the two-state model beyond −4.3. (C) Class accounting as proportional bars: of the 16,858 evaluable genes, 15,612 are undetermined; the 1,246 classed genes are 186 refractory (15 percent of the classed genes, 1.1 percent of all), 329 clustered and 731 telegraph, the last two forming the 1,060-gene two-state class. (D) Likelihood over chain depth for the twenty genes of the depth comparison, as the difference in $2 \log L$ from the two-state model (depth 1). Red: the twelve refractory genes, profiled from depth 1 to 8 (eleven extended to depth 32); the likelihood rises with depth for all twelve (8.2 to 78.6 at depth 8) and is flat beyond depth 8 for nine of the eleven extended genes and beyond depth 16 for one, and still rising at depth 32 for one, so fitted depths are lower bounds. Blue and green: the eight two-state genes of the depth comparison (seven telegraph, one clustered) and twelve more clustered genes spanning the class, fitted at depths 1, 2, 3 and 5; every chain fits worse than the two-state model for all twenty at every depth (−0.2 to −34.9 at depth 5). The clustered curves fall faster at the median (depth-5 deficit 12.1 against 7.4 for telegraph genes) without separating (one-sided Mann-Whitney $p = 0.07$). (E) Mean new-RNA count per cell for every evaluable gene, by class (violins on a logarithmic axis, black bar the median; $n = 186$, 329, 731 and 15,612; the 329 clustered genes are the two-state genes that pass the clustered test and are shown separately from the 731 telegraph genes). The refractory median (0.22) is 6.2-fold below the telegraph median (1.36); clustered and telegraph genes have the same expression level ($p = 0.68$). The axis is logarithmic because the means span five decades.

genes, and the median silent period is 95 hours against 3.9 hours for a two-state gene (Mann-Whitney $p = 9 \times 10^{-97}$). Consequently, most refractory genes are silent for longer than a cell cycle: 92 percent (171 of 186) for longer than a 20-hour fibroblast cycle. Furthermore, the active period is three-fold longer (5.8 against 1.9 hours), the burst size about two-fold larger (9.3 against 4.4 transcripts), and the message four-fold more stable (mRNA half-life 13 against 3.1 hours). The synthesis rate while active is the weakest of the contrasts, with class medians within about 20 percent (0.79-fold in the refractory class, $p = 1.3 \times 10^{-5}$) and within 15 percent after expression matching (0.86-fold, $p = 0.011$). The clustered genes are similar to the telegraph class on every one of these axes but two (Figure 3A–D). For the 329 clustered genes the silent period has a fast component with a median exit rate of 1.05 per hour (dwell time 0.95 hours) and a slow component with a median exit rate of 0.18 per hour (dwell time 5.5 hours), and it is assumed that the gene enters each of the two with equal probability after each burst. The equal weight and the fixed shape of the silence are conventions rather than fitted values; the Methods report how the clustered-class call moves when either is varied (the shape barely matters, and a heavier fast component would enlarge the class).

Clustered genes match the 731 telegraph genes in silent period (medians 3.2 and 3.3 hours, Mann-Whitney $p = 0.85$), burst size (3.7 and 3.6 transcripts, $p = 0.41$), mRNA half-life (3.2 and 3.1 hours, $p = 0.52$) and mean expression (3.3 and 3.0 transcripts per cell, $p = 0.92$). They deliver the same burst faster: their active period is shorter (1.1 against 1.8 hours, $p = 2.4 \times 10^{-24}$) and their synthesis rate higher (3.1 against 2.0 per hour, $p = 5.1 \times 10^{-32}$). The shorter active period at the same burst size makes the clustered genes noisier and they have a median Fano factor of 3.0 against 2.1 for the telegraph genes (Mann-Whitney $p = 2 \times 10^{-39}$) (Methods). Both classes differ from the refractory class on every axis: against the clustered class, refractory genes have a 29-fold longer silent period (median 95 hours), a 5.3-fold longer active period (5.8 hours), 2.5-fold larger bursts (9.3 transcripts), a 4.1-fold longer half-life (13 hours) and 1.9-fold lower mean expression (1.8 transcripts per cell), all at p below 10^{-10}. A refractory gene is therefore a promoter that is on for a few hours in every four days and, when on, transcribes at the ordinary rate. Expressed as the ratio of the silent period's half-life relative to the mRNA half-life, the two classes sit on opposite sides of a natural boundary, with per-gene medians of 5.02 for the refractory class and 0.84 for the two-state class (Figure 2E): a refractory gene's silences outlast its own message several times over, so its transcript pool empties and refills across every cycle, while a two-state gene's message integrates over several bursts.

Promoter methylation and histone modifications determine kinetic parameters

We next investigated what genetic and epigenetic features of the promoter determine promoter class and kinetic parameters. We scanned 1,019 transcription-factor motifs across the refractory promoters and an expression-matched set of two-state promoters, with per-promoter thresholds from dinucleotide-preserving shuffles and significance by label permutation within matched pairs (Methods), and found no motifs that distinguishes the refractory class beyond its composition. The only significant hit is a CREB-type site enriched for the two-state class ($q = 0.034$).

As sequence motifs alone are poor predictors of which transcription factor will bind in vivo, we also investigated evidence from ChIP-seq data. We first used Lisa[38] to compare the refractory and the two-state promoters, and this identified BRD4 and MYC among the top-ranked factors, while MED1 and SMARCA4 ranked highly for the refractory list. These factors imply RNAPII pausing as a plausible mechanism for the long refractory periods, since BRD4 recruits the pause-release kinase P-TEFb[39,40], MYC acts largely through pause release[41], and pause release has been identified as a control point of burst initiation[42]. However, measured lifetimes of paused polymerase range from a median of 16.6 seconds in the most recent genome-wide kinetic analysis[43] to medians of 7 to 17 minutes in triptolide-chase measurements[44,45], against a median refractory silence of 95 hours. We then compared the refractory promoters with the expression-matched two-state promoters using public ChIP-seq data for BRD4, MED1, MYC, and RNA polymerase II from mouse fibroblasts (Methods). Surprisingly, every factor is

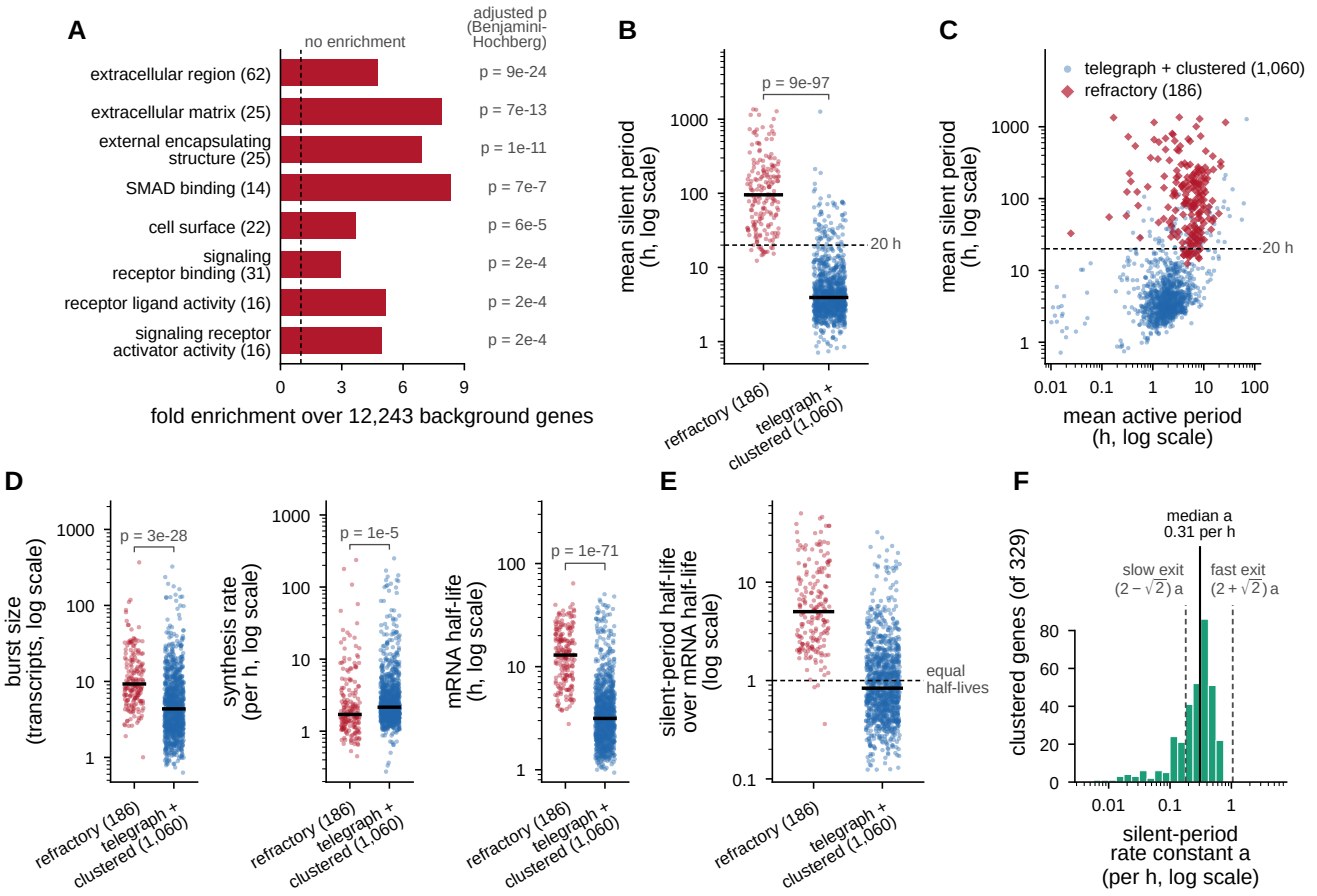

Figure 2: The refractory class is enriched for secreted and matrix proteins and is silent for days. (A) Gene Ontology cellular-component and molecular-function enrichment of the 186 refractory genes (cell-cycle-annotated genes removed; terms with at least ten refractory genes and at least two-fold enrichment over 12,243 background genes; Benjamini-Hochberg p at right). (B) Mean silent period per gene by class; the dashed line marks the 20-hour fibroblast cell cycle (171 of 186 refractory genes lie above it). Medians 95 and 3.9 hours. (C) Mean silent period against mean active period per gene. (D) Burst size, synthesis rate while active, and mRNA half-life by class, one panel per quantity; two-sided Mann-Whitney p above each pair. (E) Ratio of the silent period's half-life to the mRNA half-life; medians 5.02 (refractory) and 0.84 (two-state). (F) Histogram of the silent-period rate constant a over the 329 clustered genes under the clustered law (logarithmic bins; black line, the median, 0.31 per hour), with the slow and fast exit rates $(2 \mp \sqrt{2})\, a$ at the median marked (0.18 and 1.05 per hour); the weights (1/2 each) and the rate ratio are fixed by the law, not fitted, so every clustered gene spends 85 percent of its silent time in the slow component.

lower at refractory promoters in absolute occupancy (2.6- to 3.5-fold, against a 4.7-fold difference in fitted duty cycle), and the levels relative to polymerase show that refractory promoters are not BRD4-poor: the BRD4-to-polymerase ratio is higher at refractory promoters (median contrast +0.11, $p = 4 \times 10^{-4}$), as is the MED1 ratio (+0.13, $p = 0.015$), while MYC tracks the polymerase exactly ($p = 0.94$). Finally the pausing index of polymerase, the ratio of promoter-proximal to gene-body density, is lower at refractory promoters, not higher ($q = 2 \times 10^{-4}$; the same in the 142 pairs with the index defined on both sides, $p < 10^{-4}$). A similar result is obtained from run-on sequencing of engaged polymerase in fibroblasts where two independent PRO-seq datasets from untreated mouse embryonic fibroblast lines[46,47] have a pausing index of 1.0 in refractory promoters against 1.6 for their expression-matched two-state partners (lower in 75 and 79 percent of pairs, $p < 10^{-10}$ in each), with about half the promoter-proximal nascent signal and an unchanged 5′-to-3′ profile along the gene body (Methods). Taken together, this argues against polymerase pausing as the cause of the long inactive periods.

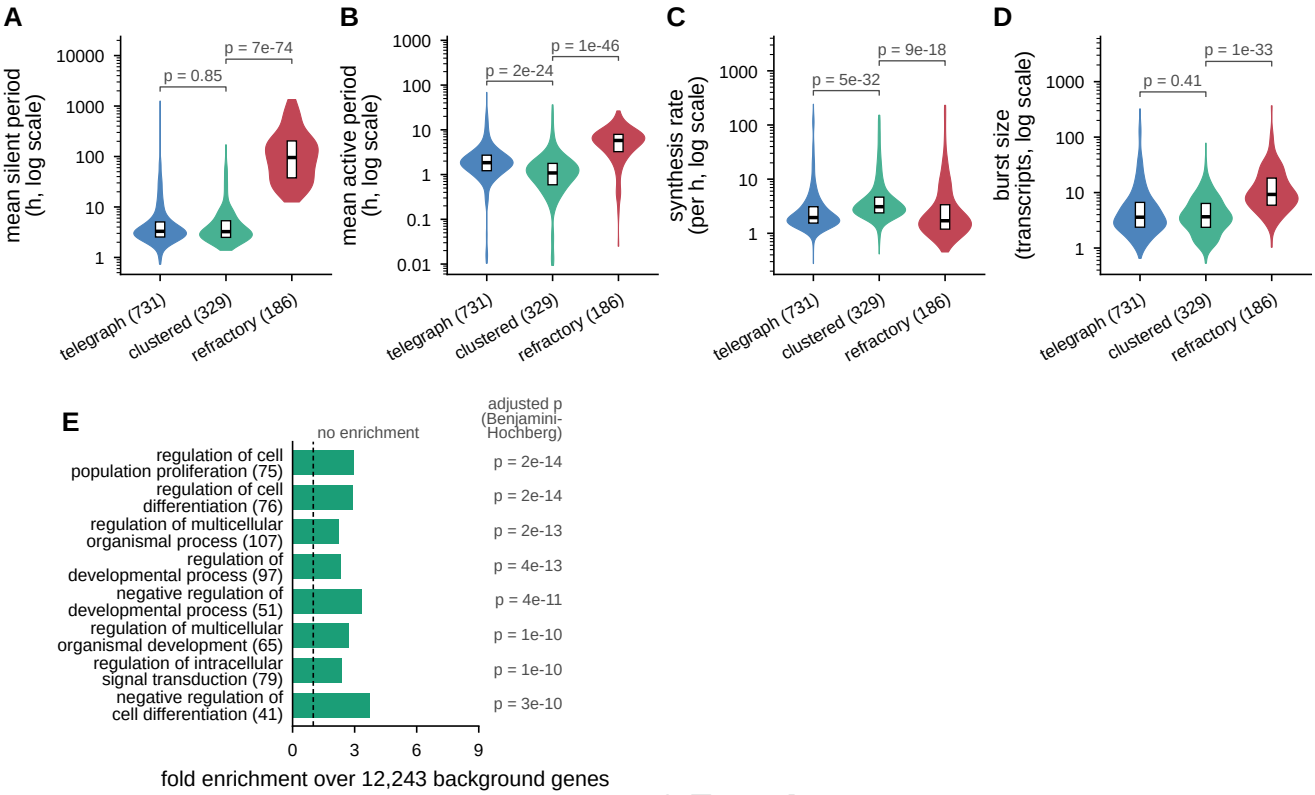

Figure 3: The clustered class matches telegraph genes except for a shorter, faster burst, and carries a developmental-regulation signature. (A–D) Mean silent period, mean active period, synthesis rate while active, and burst size for the three classes, each gene under its own class law (two-state for the 731 telegraph genes, clustered for the 329, three-state for the 186 refractory genes); black bar the median, whiskers the interquartile range, two-sided Mann-Whitney p above each panel. Clustered and telegraph genes share silent period and burst size and differ in active period and synthesis rate; the refractory class differs from both on every axis. Axes are logarithmic because the rates span two to three decades. (E) Gene Ontology enrichment of the 329 clustered genes (cell-cycle genes removed; 299 in the study set): the eight most significant biological-process terms with at least two-fold enrichment and at least ten genes, on the fold axis of Figure 2A. The clustered signature is developmental regulation where the refractory signature of Figure 2A is extracellular and secreted.

Next, we processed publicly available mouse embryonic fibroblast ChIP-seq, CUT&RUN, ATAC and DNase experiments corresponding to 163 TFs, 38 histone marks, RNAPII, ATAC, and DNAse, and we aggregated the signal within 1 kb of the transcription start site. In combination with the CpG content and the methylation data a Lasso regression model keeps 119 features to achieve $R^2 = 0.23$ for predicting the burst size (the promoter features alone give 0.19, so the added datasets add 0.04). This is substantially better than sequence features alone which only achieve $R^2 = 0.047$ for explaining burst size[6]. The largest weights are H3K27ac (+0.12), gene length (+0.09) and H3K4me3 (+0.05), with TEAD1 (+0.04) the largest transcription-factor weight. Notably, no transcription factor is above 0.05 and most of the predictive power comes from H3K27ac and gene length. Compared with their expression-matched telegraph partners, refractory promoters differ at 77 antigens ($q < 0.05$; 53 transcription factors, 21 histone marks, RNA polymerase II, ATAC and DNase), and of these 76 are lower at the refractory promoters (Figure 4A). Only H3K27me3 is higher at the refractory genes (present at 53 against 35 percent of them).

We also considered the rate at which the promoter switches to the active state. With the promoter features alone the cross-validated R^2 for the burst-on rate is 0.42, the histone marks raise it to 0.48, and the transcription factors to 0.49. A Lasso regression model keeps 168 features to achieve $R^2 = 0.49$.

The largest weights are H3K9ac (+0.26), H3K4me2 (−0.19), CpG density (+0.15), H3K4me3 (+0.13), exon count (+0.11), H3K36me3 (+0.08) and H3K27me3 (−0.08); no transcription factor is above 0.05, and most of the predictive power comes from the histone marks. Restricted to the 1,199 high-confidence genes with complete features, the pattern holds: promoter features predict the burst-on rate with 0.44 and burst size with 0.16. Across the refractory class we find that the fitted silent period increases with promoter methylation (Spearman $\rho = +0.39$, $n = 104$ genes with measured methylation, $p = 4 \times 10^{-5}$; +0.35 after controlling for expression), and decreases with promoter CpG density ($\rho = -0.36$, $n = 186$, $p = 7 \times 10^{-7}$), while the active period is uncorrelated ($\rho = -0.13$ and -0.07, both not significant), suggesting that CpG density and methylation affects only the silent state (Figure 4C and D). The same two relations hold more weakly in the two-state class ($\rho = +0.23$, $n = 508$, $p = 9 \times 10^{-8}$, and -0.30, $n = 1{,}060$, $p = 3 \times 10^{-23}$).

We also investigated the role of enhancers as it has been proposed by several groups that they can influence both burst size and frequency[6,48]. Using enhancer annotation from the literature, we carried out a similar regression analysis as for the promoters, measuring the same features at the ENCODE distal elements found within 100 kb of each promoter (219 features). The results show that enhancer features only explain 0.08 of the burst frequency and 0.04 of the burst size variance. Combining the enhancer features with the promoter features gives no improvement. The features with the highest coefficients at enhancers are histone modifications: H3K36me3 over the linked elements (+0.37), H3K27me3 (−0.29) and H3K4me2 (+0.09), with H3K27ac positively correlated but not retained by the sparse model ($\rho =$ 0.13). As for the promoters, none of the transcription factors had a coefficient greater than 0.03. Notably, several features related to 3D structure were informative: Hi-C contact density (+0.15), insulation (−0.13), and compartment (+0.10). We also quantified enhancer RNAs genome-wide (reads at intergenic distal elements within 100 kb), but they explain only 0.01 or less of either rate, and added to the promoter model they contribute nothing to burst frequency and +0.006 to burst size.

In sum, refractory promoters differ from the two-state genes in composition as evidenced by bulk assays (CpG-poor, more methylated, and lower in most active marks). This conclusion is supported by single-cell chromatin accessibility: in scATAC-seq of primary embryonic fibroblasts[49] a refractory promoter is accessible in fewer cells than its expression-matched two-state partner (median 0.11 against 0.17 of 12,124 cells with a fragment within 1 kb of the TSS, $p = 5 \times 10^{-14}$). In fibroblasts profiled for RNA and accessibility in the same cell[50], cells with neither a transcript nor an accessible promoter are more frequent for refractory genes than for two-state genes of the same expression (0.62 against 0.51 of 1,935 cells, $p = 10^{-6}$), as the long silent state predicts (Methods). Interestingly, the dependency of the kinetic parameters on these genetic and epigenetic features is largely shared. The GC content has a higher impact on k_{on} for refractory promoters, otherwise there is no statistically significant difference for the coefficients. Given these difference in DNA methylation and histone marks, we investigated the levels of the enzymes involved in adding and removing these modifications. We obtained ChIP-seq data for these proteins from publicly available datasets, and even though most are from mouse embryonic stem cells they were still included. Importantly, methylation and histone marks at promoters are consistent between mouse embryonic fibroblasts and mouse embryonic stem cells (Spearman 0.76 for H3K4me3, 0.62 for H3K27ac, 0.60 for methylation, and 0.39 for H3K27me3 over the high confidence genes), and the refractory-versus-two-state and refractory-versus-repressed differences below reproduce in the stem cells (Methods). We proceeded to investigate the activity of enzymes involved reading, writing, and erasing these modifications, and this revealed that at expression-matched pairs TET1, TET2 and TET3, DNMT3B, and the unmethylated-CpG readers KDM2A, KDM2B and CXXC1 are all found less often at refractory promoters (peaks in 65 against 90 percent of promoters for TET1, 10 against 44 percent for DNMT3B, $q < 10^{-7}$; Figure 4B). This suggests that the long silences of refractory genes are not tied to the TET–DNMT3 turnover system of CpG islands. Comparing refractory genes to fully repressed genes with matched CpG content revealed similar level of methylation (35 against 45 percent, not significant) but with repressed genes carrying twice the H3K9me3 and H3K27me3 and

a quarter of the H3K27ac (H3K27me3 and H3K27ac in Figure 4A), suggesting that the refractory promoter is methylated as its CpG content dictates but not heterochromatinized (Methods). Similarly, the histone-modifying enzymes follow the histone marks rather than the methylation: acetyltransferases (p300, CBP), deacetylases (HDAC1–3), H3K4 methyltransferases (MLL1, MLL2, MLL4, SETD1A) and demethylases (LSD1, KDM5A), the H3K27 demethylases KDM6A/B, BRD4 and BRG1 are all found less often at refractory than at expression-matched two-state promoters and more often than at CpG-matched repressed ones (for example p300 at 97, 85 and 69 percent of promoters; HDAC1 at 93, 58 and 41; Figure 4B), while the Polycomb core (EZH2, SUZ12, EED, RING1B) and the H3K9 methyltransferases (SETDB1, G9a) do not differ between refractory and CpG-matched repressed promoters.

Refractory alleles share a slow state within a gene but not across genes

Since the NASC-seq2 data allows us to resolve the chromosome of origin, we can investigate the degree of correlation between the two alleles. The two copies of the same refractory gene are more correlated with each other than the two copies of a two-state gene, once expression is matched (inter-allele overdispersion 0.59 against 0.73 over 156 expression-matched pairs, where lower means more correlated; Mann–Whitney $p = 3 \times 10^{-7}$; Figure 5B); the gap is present in every expression tercile (low -0.094, mid -0.176, high -0.067; $p \leq 0.01$), so it is not carried by residual expression differences inside the matched pairs. A direct estimate of the correlation between the two alleles' transcriptional activity in the window, from a bivariate Poisson-Beta mixture model (Methods), gives medians of 0.15 to 0.3 (paired Wilcoxon, same gene against a different gene, $p < 10^{-18}$ in both classes) for the two copies of one gene in either class, 0.02 to 0.03 for copies of two different expression-matched genes in the same cells. The shared component of allelic activity is specific to the gene rather than to the cell in both classes. The most highly correlated allele pairs belong to genes driven by a strong cell-level input: cell-cycle-periodic genes (Cdc20, Ccnb1, Ccnb2, Pclaf, Hjurp; median correlation 0.52 against 0.15 for all other classed genes) and immediate-early, interferon-, sterol-, calcium- and TGF/activin-responsive genes (Egr1, Isg15, Hmgcs1, Rcan1, Inhbb, Ski, Epha2), as expected when both copies follow the same signal.

Previous studies have indicated that nearby genes on the same chromosome are co-expressed, first in bulk expression data from yeast and human tissues[51–54] and later in single cells[55]. For every pair of autosomal genes on the same chromosome (12,557 genes with at least 50 positive cells on each allele) we compared the correlation across cells of the two genes' new-RNA indicators on the same parental copy with that of the same two genes on opposite copies. Pairs on different chromosomes give a difference of zero. On the same chromosome the same-copy correlation is the larger, and its excess decays with distance, from 8 percent of the opposite-copy value for promoters within 10 kilobases to 2 percent beyond a megabase (Figure 5C), once cells with a whole-chromosome allelic imbalance are set aside (595 cells, 6.7 percent, carry at least one chromosome whose allele-assigned reads are more than 80 percent from one parent, as expected of aneuploidy or loss of heterozygosity; chromosome-wide allelic skew is how allele-resolved single-cell RNA-seq detects both[32,56], and chromosomal instability of this kind arises in mouse embryonic fibroblasts under standard culture[57,58], they account for most of a distance-independent same-copy excess that is otherwise present chromosome-wide). Head-to-head pairs within 10 kilobases, which cannot share a transcript, show the same excess as all pairs at that distance. Inspection of the pairs with the largest same-copy excess in the genome revealed that they correspond to imprinted domains whose genes are transcribed from the same parental chromosome by construction: *Meg3–Rian* (0.23 in correlation units against a genome median of 0.001), *Dio3os–Dio3* (0.18) and *Meg3–Mirg* (0.14) in the maternally expressed *Dlk1–Dio3* domain[59], and *Ndn–Peg12* (0.14) in the paternally expressed Prader–Willi domain[60], a positive control that the statistic recovers without being told. Membership of the same contact domain in fibroblast Hi-C (L929 cells, insulation boundaries at 25-kilobase resolution) adds nothing detectable at matched distance. The fitted kinetics show a weak similarity as well: the burst-on rates of neighbours within 50 kilobases correlate at $\rho = 0.06$ ($n = 14{,}276$ pairs, Spearman $p = 2 \times 10^{-14}$; permutation over gene labels $p < 10^{-4}$), falling to zero beyond 5 megabases, burst sizes

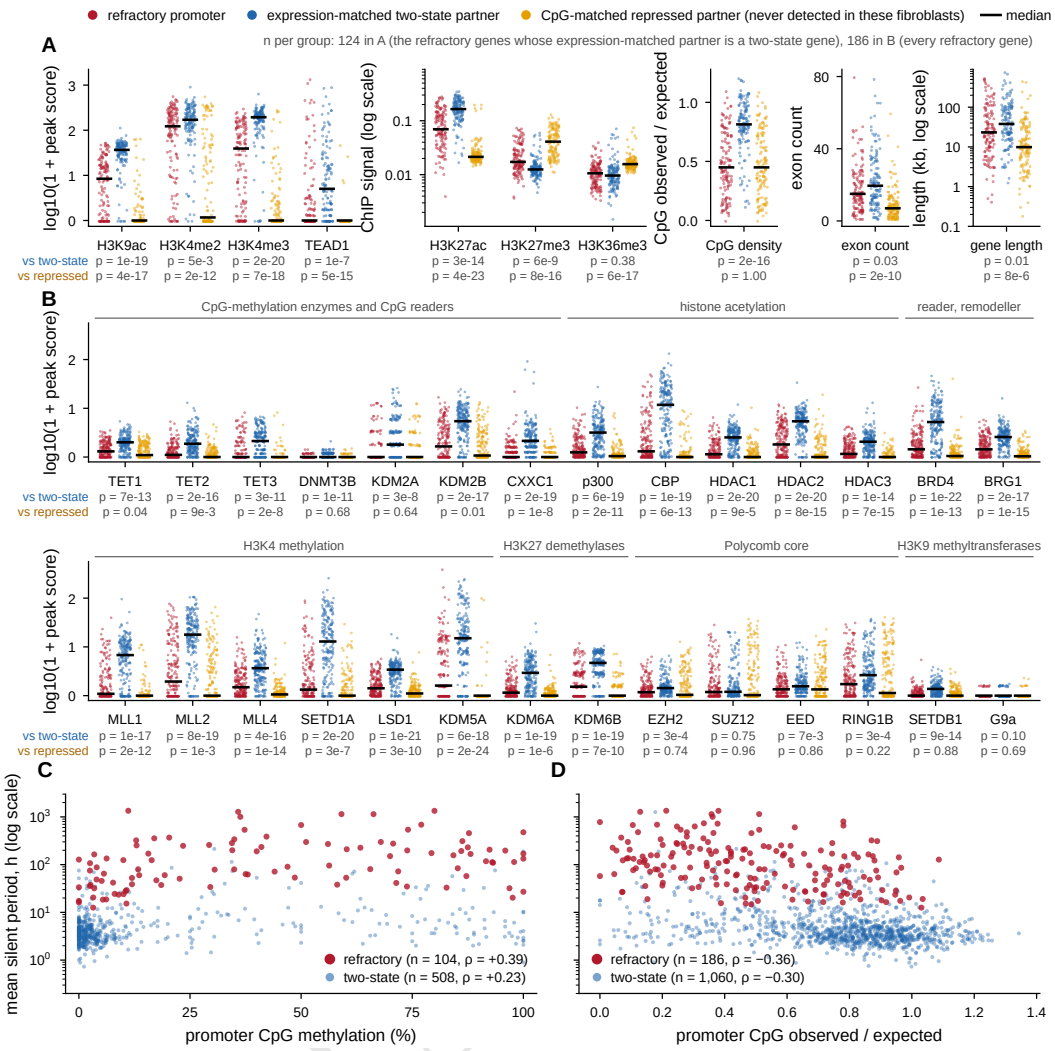

Figure 4: Refractory promoters score lower than expression-matched two-state promoters on every kinetic predictor but two (H3K27me3 is higher, H3K36me3 not different) and on every chromatin-modifying enzyme but two, score higher than CpG-matched repressed promoters on every active mark and on 20 of the 28 enzymes, and their silences lengthen with promoter methylation. (A) The promoter features with the largest weights in the burst-on-rate and burst-size regressions at refractory promoters (red), their expression-matched two-state partners (blue; 124 pairs) and CpG-matched repressed genes (orange; $n = 124$; never detected in these cells, matched on promoter CpG content; Methods): one point per gene, black bar the median; under each feature two two-sided Mann-Whitney p values, refractory against two-state (upper line, keyed in blue at the left) and refractory against repressed (lower line, keyed in orange). H3K9ac, H3K4me2, H3K4me3, TEAD1: ChIP-Atlas fibroblast peak score ($\log_{10}(1 + \text{score})$, strongest peak within 1 kb of the TSS, mean over experiments; 0, no peak); H3K27ac, H3K27me3, H3K36me3: mean ENCODE fibroblast signal over the same window (logarithmic axis, over two decades); CpG observed/expected over TSS −300 to +100 bp (equal in the repressed set by construction); exon count and gene length (logarithmic axis) from the annotation. (B) Occupancy of the enzymes named in the text (ChIP-Atlas target-gene tables pooled over all cell types, mostly embryonic stem cells; the score of A) at all 186 refractory promoters, their expression-matched two-state partners and their CpG-matched repressed genes (186 in each group); under each enzyme the Wilcoxon signed-rank p on the expression-matched pairs (upper line) and the two-sided Mann-Whitney p of refractory against repressed (lower line); points at zero, no peak in any experiment (TET1: 35, 10 and 48 percent of refractory, two-state and repressed promoters). (C, D) Mean silent period against promoter CpG methylation (MEF whole-genome bisulfite sequencing, TSS ± 500 bp; refractory $n = 104$, Spearman $\rho = +0.39$; two-state $n = 508$, $\rho = +0.23$) and against promoter CpG observed/expected (refractory $\rho = -0.36$, two-state $\rho = -0.30$); silent periods span three decades (logarithmic axis).

do not ($|\rho| < 0.05$). Correlated transcription of linked genes in single cells has been seen before, at several scales. Raj and colleagues found that activation of an integrated reporter switched on every gene in its wider locus[3], and Levesque and Raj, imaging twenty genes on single copies of chromosome 19, found a pair 14.3 megabases apart that was anti-correlated on the same chromosome copy and only weakly correlated across copies[61]. With the contrast used here, same-copy against opposite-copy correlation in allele-resolved counts from clonal hybrid fibroblasts[32], Sun and Zhang found an excess that decays with genomic distance yet stays positive beyond 60 megabases, which they attributed to shared chromatin accessibility[55]. We find that the chromosome-wide part of the excess is due to the 6.7 percent of cells with a whole-chromosome allelic imbalance, and the remainder decays over about a megabase. However, imaging of nascent transcription suggests that the coupling is due to physical rather than genomic distance: on individual human chromosomes genes co-burst when their promoters lie within about 400 nanometres[62], cohesin depletion raises co-bursting of genes tens of megabases apart on the same chromosome[63], and in yeast the divergent GAL1 and GAL10 promoters burst together on the same chromosome and not on different ones, a coupling that is lost when topoisomerases are degraded[64].

The refractory architecture detects a change in its input faster than a memoryless switch, while the two differ little in stationary output noise and burst-train information rate

To understand the functional benefits of the refractory promoters we constructed a mathematical model, evaluating three aspects. The first candidate was output noise: a silence of fixed duration should smooth the supply of transcripts relative to an exponentially distributed silence of the same mean. Computing the stationary moments of mRNA and protein exactly for both architectures, we find the effect real but small: at matched duty cycle and without feedback, across the fitted parameter range the multi-step architecture reduces the protein coefficient of variation by 4 to 13 percent depending on the number of refractory states. A multi-step silence was already known to lower output noise[22,37,65], our results provide a quantification for the high confidence genes. The second was the entropy rate of the promoter-state point process at dwell-matched parameters (Methods), the rate at which the burst train generates information about the trajectory of its own promoter. Mutual-information rates of two-state and semi-Markov promoters have been computed in closed form, and there the silent intervals did not contribute to the information exchanged[66,67]. For mouse fibroblasts, dwell-matched parameters the information rate is 0.5536 nats per unit time for the multi-step promoter and 0.5409 for the two-state promoter, and across the genome the classes are indistinguishable on this measure once dwell times are matched.

The third candidate is sensing speed, that is the time to detect a change in the input. Whereas the information rate is a property of the stationary distribution, the sensing speed is the delay in noticing a change in the input, and the two are not necessarily correlated. A cell that must notice a change in its environment reads the change from the timing of its own bursts, and the best that any cellular mechanism reading those bursts could do is the optimal sequential detector of Shiryaev[68], which represents an idealized observer of the burst train and we calibrate false alarms with its Shiryaev-Roberts form[69] (Methods). We modeled bursts as a renewal process with exponential intervals (the two-state promoter) or Erlang-k intervals of the same mean (a k-step refractory promoter), imposed a step change in the mean interburst time, and computed the expected detection delay at matched mean time to false alarm (Figure 7). To evaluate the differences in detection time we considered measurements for the estrogen-responsive TFF1 gene[23] (interburst time 185 minutes at the half-maximal dose, 66 minutes at saturation). A depth-seven refractory promoter detects the change 3.9 times sooner than the exponential promoter at one false-alarm calibration (444 minutes reduced to 114) and 4.7 times sooner at a stricter one (803 reduced to 170); the second reported saturating value, 86 minutes, gives 3.7-fold. These delays are conditional on detection occurring after the change, which under the simulated change-time prior holds for about 5 percent of realizations at the first calibration and 35 percent at the second. The

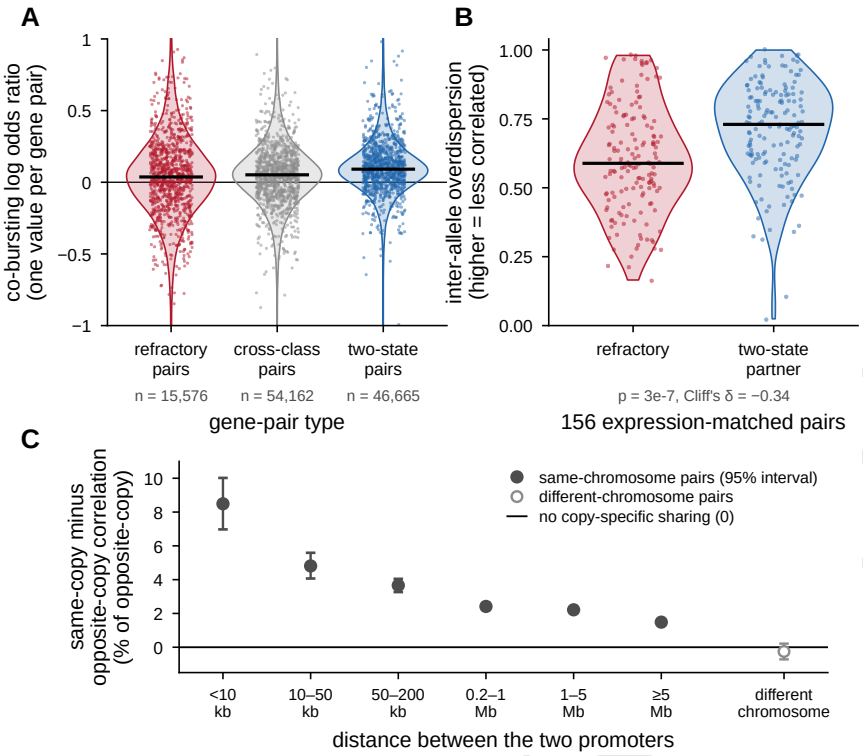

Figure 5: Refractory alleles share a slow state within a gene but not across genes: gene pairs of the refractory class show no excess co-bursting, while the two alleles of a refractory gene are more correlated than those of matched two-state genes. (A) Co-bursting of gene pairs, one value per pair: the Mantel–Haenszel log odds ratio of the two genes' bursting indicators across cells, stratified by decile of the cell's total molecule count (Methods), for refractory pairs (red; $n = 15,576$), cross-class pairs (grey; 54,162) and two-state pairs (blue; 46,665). Violins are the densities of all pairs at equal maximum width, points a random sample of 1,000 pairs per group, black bars the medians (0.037, 0.052, 0.091, in the order of the class duty cycles); the axis omits the fewer than 1 percent of pairs beyond ± 1. The refractory-minus-cross-class contrast, -0.015, lies below the null obtained by permuting each gene's bursting indicator across cells within depth deciles (99th percentile $+0.0014$; one-sided permutation $p = 1$ for an excess): no class-specific co-bursting. (B) Inter-allele overdispersion of each refractory gene (red) and its expression-matched two-state partner (blue), 156 pairs, violins of the two distributions behind the points (higher means less allele correlation; axis from zero); black bars the medians, 0.589 against 0.730; Mann–Whitney $p = 3 \times 10^{-7}$, Cliff's $\delta = -0.34$. (C) Co-transcription of neighbouring genes is copy-specific and decays with distance: same-copy minus opposite-copy correlation of the new-RNA indicators of gene pairs on the same chromosome, as a percentage of the opposite-copy correlation, by distance between promoters (1,776 pairs within 10 kb to 4,054,339 beyond 5 Mb; cells with a whole-chromosome allelic imbalance excluded; bars are 95% intervals), with pairs on different chromosomes as the control (open marker, zero within its interval).

advantage belongs to regular silences and reverses for irregular ones: TFF1 itself bursts in clusters, with silences more variable than exponential[23], and a clustered promoter with squared coefficient of variation 2 at the same mean interval detects the same change 1.3 to 1.5 times later than the memoryless promoter and 5 to 7 times later than the depth-seven refractory promoter (Methods), because its intervals carry 0.62 of the memoryless information about the change. This holds for a dose that rescales the silences; a dose acting on the short silences alone, leaving a long dormant state untouched, is a different comparison in which the mean interval barely moves (Methods).

Detection delay is non-increasing in chain depth across the tested grid, and the advantage grows as the detection problem gets harder. By a classical result, the expected delay of an optimal detector scales inversely with the information each interval carries about the change, the Kullback-Leibler divergence

between the pre- and post-change interval distributions[70]. For Erlang intervals of fixed mean that divergence is k times the exponential value[71], so a regular clock reveals a rate change several-fold faster than a memoryless one. Regular baselines make rate changes easier to detect, a principle known from sensory spike trains[72], and kinetic models had predicted that refractory or multi-step promoters respond to a stimulus faster[73,74] and with higher specificity[75]. That is, a refractory promoter is slower to fire after any single activating event, but at matched mean interval it is faster to reveal that the input rate has changed, because the detector uses the regularity of the intervals. Against a memoryless two-state promoter of the same mean interval, the five-step refractory promoter detects the change sooner at every one of the 48 settings tested (Methods): when the mean interval changes 2.8- or 24-fold the two-state promoter's delay is 2.0 to 3.8 times the refractory promoter's (median 3.1), and for the smaller changes of 1.01- to 2.7-fold that a dose acting on a single kinetic component produces it is 1.02 to 3.7 times (median 2.4). With two silent steps instead of five the factors are 1.6 to 1.8 and 1.0 to 1.75. The refractory promoter's advantage over a two-state one therefore holds for any change in the mean interval and grows with the number of steps and the size of the change.

Both quantities have closed forms, so we can ask directly whether noise and sensing speed trade off (Methods). The Fano factor of the mRNA count depends on the silent period only through the mean of $e^{-k_\mathrm{d}T_\mathrm{off}}$ over the distribution of the silent period, which we write with angle brackets as $\langle e^{-k_\mathrm{d}T_\mathrm{off}}\rangle$:

$$F = 1 + \frac{k_\mathrm{syn}\, k_\mathrm{off}\, (k_\mathrm{d} - k_\mathrm{on}\varphi)}{k_\mathrm{d}\, (k_\mathrm{on} + k_\mathrm{off})\, (k_\mathrm{d} + k_\mathrm{off}\varphi)}, \qquad \varphi = 1 - \langle e^{-k_\mathrm{d}T_\mathrm{off}}\rangle,$$

which reduces to the two-state expression $F = 1+k_\mathrm{syn}k_\mathrm{off}/((k_\mathrm{on}+k_\mathrm{off})(k_\mathrm{on}+k_\mathrm{off}+k_\mathrm{d}))$ for an exponential silence. A more regular silence lowers F below its two-state value: at the fitted rates of the 1,246 high confidence genes the median falls from 2.51 for the one-step (two-state) silence to 2.28 for two steps and 2.09 for seven. The delay in detecting a step in the burst-on rate from k_on to k'_on also falls with depth. When bursts are resolved, the delay at a mean time to false alarm γ is $\ln\gamma/I$ to first order[70], with $I = k[\ln(k'_\mathrm{on}/k_\mathrm{on}) + k_\mathrm{on}/k'_\mathrm{on} - 1]/(1/k'_\mathrm{on} + 1/k_\mathrm{off})$ for a k-step silence. Our simulations show that the delay is $1/k$ of the memoryless value at a large false-alarm budget, and 0.65 (two steps) or 0.50 (three) at a budget of 100 promoter cycles. Depth therefore improves both, and the architecture itself carries no trade-off between noise and speed.

Detection speed measures the response to a single step. We also asked how much a promoter can tell the cell about an environment that keeps changing. We placed a two-state environment, switching every 1, 10, or 100 mean mRNA lifetimes ($1/k_\mathrm{d}$, the unit of time in the model), in front of each promoter and let it drive one parameter at a time (synthesis rate, burst-on rate, or active-exit rate) so that the high state raised the output mean by 1.5- or 3-fold. For each case we computed exactly the rate at which the count sequence carries information about the environment, the tracking error of the optimal filter, and the low-state Fano factor (Methods). The synthesis-rate channel carries the most information at 31 of 36 combinations of architecture, duty cycle, and operating point. The burst-frequency channel is most informative for the remaining five combinations, all at fast switching with a small mean and a small change. Between the two switching architectures at a matched duty cycle, the three-step promoter carries more information about the environment, measured as the information rate at frame boundaries, than the two-state promoter at 16 of 18 points, by a median factor of 1.6, and has a lower low-state Fano factor at all 18. We conclude that the multi-step architecture is better for detecting both types of changes. Note that this dynamic measure does not contradict the static ordering, in which added promoter states lower channel capacity[76,77]. The trade-off between noise and detection speed appears only across operating points: at the fitted class medians (duty cycles 0.052 and 0.32) the refractory operating point is the less informative and the noisier at every operating point where the two can be matched, the same split between operating point and architecture that the archetype analysis below shows. The environment enters these calculations at frame boundaries, so the information rates are lower bounds on the information carried about the full environment path.

We also asked the question from the other side: if the refractory class is suited to fast detection, do the classes form the kind of polytope in parameter space that archetype analysis would reveal[78], with a detection task at one corner? We used principal convex hull analysis[79] to fit a polytope with three to six vertices to the points represented by the four dimensional vector (k_{on}, k_{off}, k_{syn}, and degradation rate) for each of the 1,246 high confidence genes. We obtained the best fit for three archetypes, and we find that the refractory class is a good match to one of the archetypes with rare long silences, large bursts, and long-lived message. Of the hundred genes nearest it, 86 percent refractory against a 15 percent base rate ($p = 4 \times 10^{-50}$). The other two archetypes are both two-state corners with the same three-hour silence and burst size of about 3.5, separated along the active-period axis (1.9 hours at low synthesis against 0.7 hours at threefold higher synthesis). The two two-state corners have the same mean and burst size and differ in tempo: the short-ON corner has half the promoter correlation time (0.6 against 1.2 hours), nearly twice the information rate (0.83 against 0.45), half the detection delay (3.6 against 5.7 cycles; 14 against 31 hours) and a modestly higher Fano factor (3.1 against 2.1), while the refractory corner has a Fano factor of 10, an information rate of 0.11, the fewest cycles to detection (3.1) and the most hours (189). Both two-state vertices have silent periods near 3.4 hours and burst sizes near 3.5 molecules; one (V0) reaches that burst size through a long, weakly productive active period (1.9 hours at 1.8 molecules per hour, duty cycle 0.35) and the other (V2) through a short intense one (0.7 hours at 5.1 molecules per hour, duty cycle 0.17). Nothing at the promoter separates the genes nearest the two vertices: GC and CpG content, CpG-island overlap, methylation, H3K4me3, H3K27ac, H3K27me3 and H3K9me3, and 156 factors and marks profiled earlier are all indistinguishable, as are their Gene Ontology annotations. The clustered subclass of the two-state genes is spread across both vertices: it makes up 21 percent of the two-state genes among the hundred nearest V0 and 36 percent near V2 (base rate 31 percent). What separates the two archetypes lies downstream of initiation. V0 genes are longer (median 50 against 27 kilobases; Mann–Whitney $p = 3 \times 10^{-3}$) with more exons (25 against 16; $p = 2 \times 10^{-5}$), and at equal polymerase occupancy of the promoter they hold more of it near the start site by ChIP-seq (pausing index 12.8 against 7.9, $p = 1 \times 10^{-3}$), carry fewer engaged polymerases in the gene body by run-on sequencing in two fibroblast datasets while their promoter-proximal run-on signal and $5'$-to-$3'$ processivity are unchanged, and in these cells' own labelled reads their new RNA is more intronic (19 against 9 percent of new reads) and sits further $5'$ relative to the same genes' old RNA (the new-read fraction falls more steeply along the gene body near V0; Mann–Whitney $p = 0.01$, paired $p = 0.04$ over 66 expression-matched pairs) (Methods). Across all 1,060 two-state genes the fitted active period does not lengthen with gene length (Spearman -0.06 with expression held fixed), so the long active periods of V0 are not an artefact of transcript completion time, whereas synthesis rate and burst size do fall with gene length (-0.19 and -0.26). We interpret the V0–V2 axis as one of polymerase release and early elongation on long, exon-rich genes, orthogonal to the promoter-state axis that separates the refractory class.

Refractory calls on the X chromosome survive the single-copy correction that X inactivation forces, with one open sensitivity

Our methodology was developed for diploid cells, but X inactivation makes X-linked genes effectively single-copy in each cell, and in these data this should result in a bimodal per-cell X haplotype share at zero and one. Indeed, 95 percent of cells carry a dominant X haplotype above 0.9, and the ten refractory genes on the X chromosome are 93 to 98 percent monoallelic. Since our genome-wide fits, X included, used a two-allele model we refit 45 confidently classed X-linked genes under a one-copy law (Figure 9). Eight of the ten refractory promoters hold under the corrected law (Tspan7, Gpc3, Tnmd, Srpx, Rnf128, Arxes2, Tmem185a, Efnb1), but Vegfd and Hs6st2 are instead classed as two-state. Conversely, two genes in the two-state class acquire a nominal refractory statistic under the one-copy law; one is a gene that escapes X inactivation, to which the law does not apply, and the other's one-copy fit did not converge. One-copy fits met the run's convergence criterion for 68 percent of genes, against 96 percent

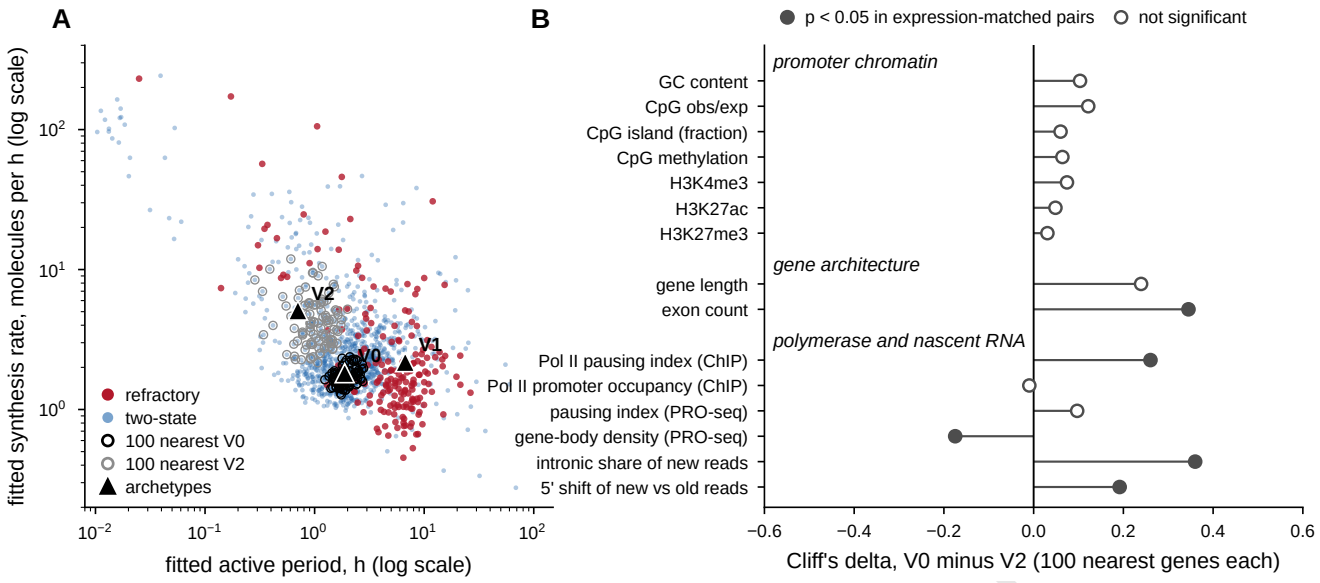

Figure 6: The two two-state vertices differ downstream of initiation, not at the promoter. (A) The 1,246 classed genes in the plane of fitted active period and synthesis rate (both logarithmic, since the vertices differ multiplicatively), refractory genes in red and two-state genes in blue, with the three archetypes of the principal convex hull fit (triangles; V1 is the refractory vertex) and the 100 two-state genes nearest V0 (black rings) and V2 (grey rings). (B) Cliff's delta, V0 minus V2, for the 100 genes nearest each vertex, for promoter composition and chromatin (top), gene architecture (middle) and polymerase and nascent-RNA measures (bottom: Pol II ChIP-seq in embryonic fibroblasts, PRO-seq in an embryonic-fibroblast line, and the labelled reads of this dataset realigned); filled markers, $p < 0.05$ in expression-matched V0–V2 pairs. Promoter features do not differ; V0 genes are longer with more exons, hold more polymerase near the promoter, carry less in the body, and accumulate new RNA that is more intronic and more 5′-proximal.

for the two-copy controls, though all ten refractory genes converged under both laws. The parameter correction the one-copy law forces is the model's own copy-number degeneracy: two alleles bursting at rate k_{on} are nearly indistinguishable from one allele bursting at $2k_{\text{on}}$, and under the corrected law the burst-on rate of every refractory gene on X doubles (ratios 1.9 to 2.5, median 1.9). We observe that the two-allele fits had absorbed the missing copy into a halved burst frequency, so the burst-on rates previously reported for X-linked refractory genes in the genome-wide table are about two-fold low.

Discussion

The telegraph model owes its standing to two facts: it is the simplest promoter with a silent state, and a snapshot of single-cell counts cannot easily reject it. Time-resolved data allows us to rigorously test this model against alternative, more complex models. Using NASC-seq2 data we demonstrate that a subset of promoters have more regular off-states than a memoryless promoter and are best fit by a multi-step, refractory model. Notably, multi-step silences have also been reported for mammalian genes from luminescence time series[5,22]. The refractory class is enriched for secreted and matrix proteins, it is associated with methylated CpG-poor promoters, it is silent for longer than a cell cycle, and its silences lengthen with promoter methylation. We note that both the gene category and the promoter state have been described before, in each case without kinetics. The secreted and matrix category is where days-long repression[23], stable allelic silencing in fibroblasts[32,33] and heritable expression states in macrophages[80] have each been reported, and we hypothesize that the refractory class constitute a subset inside that category. The promoter state corresponds to the 'low-CpG promoter' class reported by Weber et al., which in primary fibroblasts is "methylated in the inactive as well as in the active

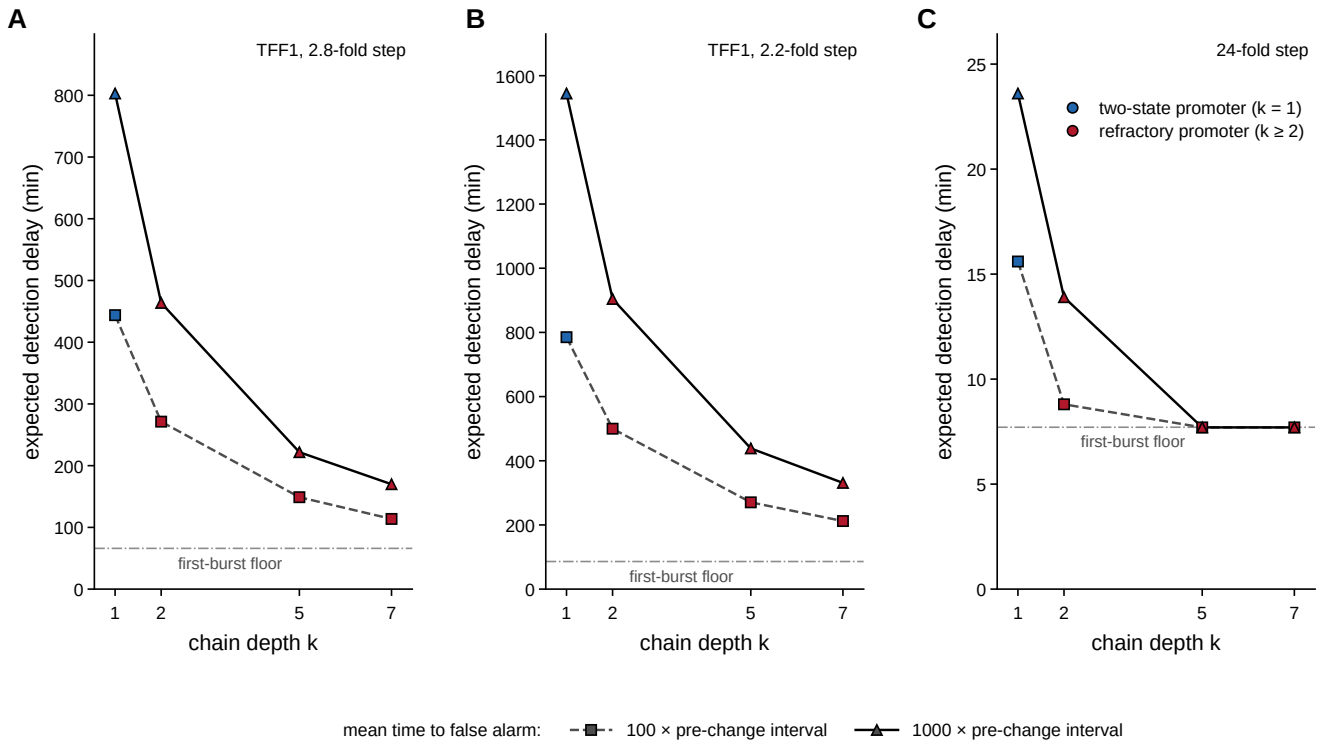

Figure 7: Detection delay falls with chain depth. Expected delay of the optimal sequential detector after a step change in burst rate, for Erlang-k silent intervals of the same mean (blue, $k = 1$, the two-state promoter; red, $k \geq 2$, refractory promoters), at two false-alarm calibrations: a mean time to false alarm of 100 (dashed, squares) or 1,000 (solid, triangles) pre-change intervals. (A) The TFF1 operating point, interburst time 185 minutes falling to 66 (a 2.8-fold step). (B) The second reported saturating value, 185 falling to 86 minutes (2.2-fold). (C) A 24-fold step at the same interval scale, 185 falling to 7.7 minutes, on a vertical scale sixty-fold finer than A and B; deep chains reach the first-burst floor (dash-dot line, the post-change mean interval), where depths five and seven tie. Delays are conditional on detection after the change, which under the simulated change-time prior holds for about 5 percent of realizations at the first calibration and 35 percent at the second.

state"[81] and essentially never carries H3K27me3[82]. Induced low-CpG promoters that stay methylated while gaining acetylation and elongating polymerase were described in primary mouse dermal fibroblasts[83], and cytosine methylation coexisting with H3K27ac has been reported at enhancers though not at promoters[84]. Those studies established that such promoters can be active while methylated, but none of them reported promoter kinetics.

We report that this class of refractory promoters is defined by the timing of its silences which are more regular than a memoryless switch and longer than a cell cycle. The length of the silence proportional to the level of promoter methylation while the active period is not. Moreover, refractory promoters carry the same machinery as their expression-matched two-state counterparts, and if anything more BRD4 and Mediator per polymerase rather than less. By contrast, recruited DNA methylation drives a locus into silence with a half-time of about 62 hours[85] and heterochromatin spreads at hours per nucleosome[86]. Thus, we conclude that these are plausible mechanisms for a promoter that must be re-opened through several such steps to produce silences with the multi-step structure we fit. The recruited state, however, does not reactivate[85,87], so a silence that ends after 95 hours must rest on methylation that turns over rather than saturates.

Interestingly, Czapiewski and colleagues, finding CpG-island promoters less noisy than CpG-poor ones in single-nucleus counts, proposed that slow chromatin-remodeling transitions generate long off-states at promoters without islands[88]. The refractory class confirms this prediction and adds that the off-states of these promoters are regular, not merely long. Moreover, the allele data are consistent with

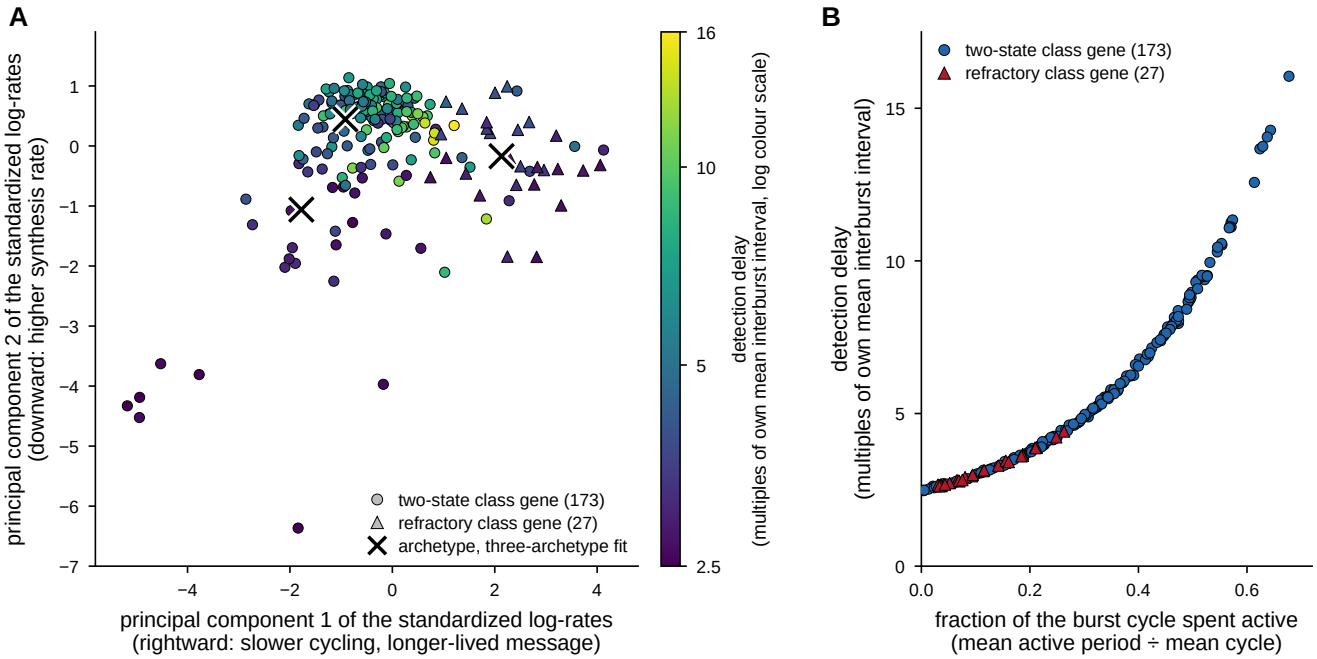

Figure 8: Across the promoter cloud, operating point rather than architecture sets detection delay. (A) Delay under the optimal sequential detector (2.8-fold change in burst rate; mean time to false alarm 100 times the promoter's own mean interburst interval) at two hundred genes spanning the cloud (173 two-state, circles; 27 refractory, triangles), projected to the top two principal components of the standardized log-rates over the 1,246 classed genes; the delay is in units of each promoter's own mean interburst interval, on a logarithmic color scale from 2.5 to 16. Crosses mark the vertices of the three-archetype fit, shown for orientation only (no archetype number gives a significant polytope). The fastest detectors are the high-synthesis two-state genes at the lower left, not the refractory corner. (B) The same two hundred delays against the fraction of the burst cycle the promoter spends active: genes of both classes lie on one curve, because every gene is evaluated with the same renewal detector at its own rates, so at this level of comparison the operating point alone sets the delay; the refractory advantage of Figure 7 is an architecture effect at a matched operating point.

such a clock, with no coordination across genes. The genes where the two copies are highly correlated shared cell-level control of activation[89,90], consistent with previous studies. Kalo et al showed that four live β-actin alleles in fibroblasts activate independently at steady state and synchronously under serum, and the excess over a binomial was fitted as two cell subpopulations[90]. Rodriguez et al found that the slow repressive state of TFF1 is occupied independently by its two alleles, with the fitted coupling acting on the activation step[23].

Our model also allows us to address coordination between the two alleles: the two copies of one gene are correlated in both classes (median correlation 0.15 to 0.3, against 0.02 to 0.03 for two different genes in the same cells), and once expression is matched the two copies of a refractory gene are the more correlated pair (Figure 5B). Allele-resolved single-cell RNA-seq of preimplantation embryos established independent allelic transcription genome-wide from pooled biallelic fractions[91], and on the present data the two alleles of a gene have been described as independent without a reported test[19], while positive inter-allele dependence at high expression has been read as the effect of diffusible factors[11,92,93]. Following the pioneering work using two-reporter experiments[3,94–96], correlation between the two copies of one gene in one cell has been interpreted as the extrinsic part of expression noise, the part set by the state of the cell rather than by the promoter. Allele-resolved single-cell RNA-seq and allele-specific imaging in mammalian cells have found the two copies bursting largely independently, with a shared part attributed to cell state, cell-cycle stage or a common stimulus[19,91,97–99]. In mouse fibroblasts the extrinsic noise has been estimated for some four thousand genes[100], and it has also been shown that

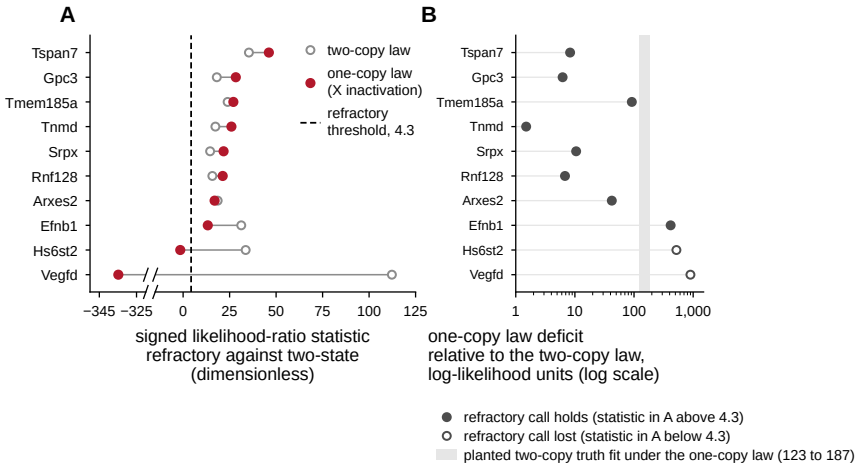

Figure 9: Refractory calls on the X chromosome under the single-copy law. (A) Signed likelihood-ratio statistic (refractory against two-state) for the ten X-linked refractory genes under the two-copy law used genome-wide (open) and under the one-copy law that X inactivation implies (filled); the dashed line is the class bar of 4.3; the axis is broken to reach Vegfd's one-copy value. Eight calls hold; Vegfd and Hs6st2 fall below the bar. (B) Log-likelihood deficit of the one-copy law relative to the two-copy law per gene (filled, call holds; open, call lost). The shaded band is the deficit a planted two-copy truth pays when fit under the one-copy law at these genes' rates (123 to 187 units); a planted one-copy truth gives the opposite sign (the one-copy law wins by 11 to 43 units), off this axis. The two lost calls sit at 520 and 900 units, far above either planted cost, so the one-copy law is rejected by their counts.

the degree of coordination between the two copies differs from gene to gene in embryos and during reprogramming[11,101,102]. The comparable single-gene numbers from previous studies vary widely, with about 0.55 for Nanog in mouse embryonic stem cells[97], 0.54 for the integrated output of the two GREB1 alleles and 0.01 for their instantaneous bursting[99], 0.05 to 0.07 for nascent counts at the two copies of Oct4 and Nanog[98], and 0.06 for two reporters at different loci[3]. Genome-wide, 2 to 6 percent of bursty genes depart from allelic independence, nearly all toward coordination[11]. The explanation for the wide range is that the lower ones are for instantaneous correlation while the higher values are integrated across a time window, and thus they are not directly comparable to our values. What is both novel and robust though is the class difference: the two copies of a refractory gene share more of their activity than the two copies of a two-state gene, and the shared part belongs to the gene, not to the cell.

We investigated what could be the benefits of the refractory architecture over the memoryless one. A secretory cell that must respond to a change in its environment reads that change from the timing of its own bursts. A regular clock reveals a rate change several-fold faster than a memoryless one, because the information per interval about the change grows linearly with the number of steps and detection delay falls inversely with that information[70,71]. By the same rule an irregular clock, such as the clustered promoter of TFF1[23], reveals it more slowly than a memoryless one. Our closed form solution shows that additional steps in the refractory chain lowers both the noise and the detection delay, and the architecture carries no trade-off between them. Across the real genes the two rise together instead: over the 1,246 classed genes at their fitted rates the Fano factor and the delay in hours for a 1.5-fold step correlate at Spearman $\rho = 0.74$, because large bursts raise the noise and long cycles lengthen the delay, and the refractory class is both the noisiest (median F 6.3 against 2.1 for two-state genes) and, in hours, the slowest (about 1,100 against 110 hours at a budget of 100 cycles), while counted in each gene's own cycles it is the fastest (22 cycles against 50). That is, the advantage of the refractory promoters for detection speed is an architecture effect at a matched operating point. Measured in absolute hours, the refractory corner of our archetype analysis is the worst for detection. The reason for this dichotomy is that a regular promoter needs fewer intervals to reveal a change, and the refractory genes' intervals are

an order of magnitude longer, which is a property of their operating point and not of their architecture. For a program of secreted proteins, whose output is committed to the extracellular space and cannot be recalled, deciding quickly and correctly when to restart is a plausible advantage.

A silent period spanning multiple cell cycles has a formal alternative, a stable subpopulation of cells in which the gene is off. The absence of any excess co-bursting across the class (Figure 5A) argues against a shared cell state, but a per-gene silent subpopulation is not excluded by snapshot data. Such a subpopulation can be of the clone-level kind in which a gene is off in both alleles of some clones[103], of the heritable kind reported for cytokine and chemokine genes over more than 25 divisions[80], or, at the level of single alleles, of the clonally stable monoallelic kind[32,104–106]. In fibroblasts that last class is itself low-expressed and enriched for secreted and extracellular proteins[32]. Whether such a heritable state sits inside the refractory class is a question that can only be settled using single-cell time series.

Although our analysis of methylation and histone data have uncovered several key properties of refractory promoters, additional experiments are required to test some of our predictions. We found that the duration of the refractory period is proportional to the DNA methylation, and thus lowering promoter methylation, by DNMT inhibition or a DNMT1 degron, should shorten refractory silences without changing their multi-step shape, or affecting the two-state class. Moreover, our mathematical model of detection timing predicts that after a change in the cell the refractory genes turn on later in hours but sooner in their own burst cycles than the two-state genes. We found evidence of this in a skin-wound time course profiled using bulk RNAseq[107] where the refractory genes had completed a median 42 percent of their day-14 expression change by day 7, against 53 percent for expression-matched two-state genes ($p = 2 \times 10^{-4}$). Finally, single-molecule imaging with metabolic pulse-chase labeling such as TEMPOmap[108] using a labeling window of comparable length to the mean off-time would provide further insights into the underlying mechanism.

Limitations

The three parameters of the Poisson-Beta model cannot be fit directly from the NASC-seq2 data as degradation rates must be supplied exogenously to fix the scale. Our assignments to the two-state and refractory classes are insensitive to scaling the degradation rate, but the absolute fitted rates are impacted. The clustered law was fitted with one fixed shape (squared coefficient of variation 2, below the 2.3 to 4.3 measured for a clustered promoter in live cells[26]), and it was infeasible to fit a free mixture weight. Only 1,246 of 16,858 evaluable genes were confidently classed as refractory, clustered, or telegraph, and the main limitation is the low number of transcripts detected. Comparisons between the classes was limited by the refractory genes, and with only 186 members this limits our power detect differential motifs. Similarly, our enhancer comparison is underpowered and the use of a 100kb window to assign enhancers to promoters likely results in a large number of false positive and negatives. The chromatin and methylation data were from other studies and thus the relationship to the inferred kinetics are correlatory, not causal. A methylation-expression correlation is equally consistent with methylation as cause or as passenger of promoter occupancy[109], and an antisense-induced promoter memory that carries methylation has not been excluded[110]. The sensing-speed result is a calculation at measured operating points, checked against reported dose-response statistics for one gene, not a trace-level measurement. The inter-allele correlation result is a contrast between classes; a class-differential counting artifact makes the absolute values uninterpretable as correlation levels, and the contrast is compatible with a shared silent state, a shared cell-level control of activation or an interallelic coupling, which windowed counts do not separate[89,90,111].

Methods

Data and software

Resource	Source	Identifier
Deposited data		
NASC-seq2 new-RNA counts, primary mouse fibroblasts (F1 hybrid), 2-h 4sU window	Ramsköld et al.[19]	ENA: PRJEB60799
Per-gene mRNA degradation rates, mouse fibroblasts	Larsson et al.[6]	Zenodo record accompanying the paper
MEF whole-genome bisulfite sequencing (CpG methylation)	public track, see Analysis details	to be listed at submission
Histone-mark ChIP-seq, mouse fibroblasts (H3K4me3, H3K27ac, H3K27me3, H3K9me3, H3K36me3)	ENCODE	accessions to be listed at submission
ChIP-seq, MED1 (MEF), MYC (MEF), RNA polymerase II (MEF), BRD4 (NIH/3T3)	Sequence Read Archive	SRX accessions listed in Table S1
GENCODE vM25 gene annotation (mm10)	GENCODE	vM25
JASPAR CORE vertebrate motifs (1,019 matrices)	JASPAR	latest version per base matrix identifier
Gene Ontology annotations, mouse	MGI GAF and go-basic OBO	downloaded 2026
TFF1 burst-kinetic operating points	Rodriguez et al.[23]	Cell 176:213
scATAC-seq, untreated primary mouse embryonic fibroblasts (129/Ola), fragment file; matched scRNA-seq	Muckenhuber et al.[49]	GEO: GSE160764 (GSM5852360, GSM5852363)
Single-cell Multiome (RNA + ATAC), Cas9 mouse embryonic fibroblast line, control-guide cells	Burr et al.[50]	GEO: GSE297418 (GSM8990892, GSM8990896)
PRO-seq, immortalized wild-type mouse embryonic fibroblasts, untreated, normalized strand-specific coverage	Mahat et al.[46]; Vihervaara et al.[47]	GEO: GSE71708 (GSM1843438); GSE128160 (GSM3665947)
Software and algorithms		
Exact transient likelihoods for the two-state and multi-step promoter models, exact p-values, control constructions	this paper[28]	code to be deposited at submission
Lisa	Qin et al.[38]	web service
Principal convex hull (archetype) analysis	Mørup and Hansen[79]	reimplemented, ten restarts
GOATOOLS, JAX, SciPy, NumPy, pyBigWig	open source	versions to be listed at submission

Correspondence and availability

Correspondence Further information and requests for resources should be directed to and will be fulfilled by the corresponding author, Martin Hemberg (mhemberg@bwh.harvard.edu).

Materials This study did not generate new reagents.

Data and code availability All data analyzed are public (Data and software table). Per-gene fitted rates, class calls, exact p-values, and the analysis code will be deposited in a public repository before publication; the accession will be listed here.

Cells and data

No new experiments were performed. The study reanalyzes published NASC-seq2 data from primary mouse embryonic fibroblasts of an F1 hybrid cross [19], together with public chromatin and ChIP-seq data from mouse embryonic fibroblasts and NIH/3T3 cells.

Analysis details

Data and labeling window We used the new-RNA count matrix for the 8,916 cells with a two-hour 4-thiouridine window, allele-resolved by the F1 hybrid single-nucleotide variants, as released with the method paper. New-RNA counts and total counts per gene and cell, and per-cell allele-specific new counts, are the inputs to every fit.

Gene inclusion The reconstructed-molecule matrices contain 37,249 genes by 8,916 4sU-treated cells. Genes were kept if at least 50 cells carried at least one molecule, new or old (13,484 genes removed, 99 of them with no molecules at all), and if the new fraction of their molecules lay between 2 and 98 percent (480 removed, 473 of them below 2 percent), leaving 23,285. The transient likelihoods are evaluated on a count grid of 1.6 times each gene's maximum count plus a margin, and 39 genes with very large count ranges (31 of them among the 1,000 most abundant, including Actb) exceeded the solver's budget and were not fitted, leaving 23,246. 16,858 genes are evaluable, meaning that both the two-state and the multi-step fits converged (gradient sup-norm below 10^{-3} or final step below 10^{-5}). 798 converged for the two-state model only, 1,562 for the multi-step model only, and 4,028 for neither. Both alleles are pooled in these fits; allele-resolved counts enter only the allele analyses.

Promoter models Both models have an active state from which transcripts are synthesized at rate k_{syn} and are degraded at rate k_{d}, and the active state is left at rate k_{off}. In the two-state model the silent state is left at rate k_{on}, so silent periods are exponential with mean $1/k_{\mathrm{on}}$. In the multi-step model of depth k the promoter returns to the active state through k sequential silent states, each left at rate $k\,k_{\mathrm{on}}$, so silent periods are Erlang-k with the same mean and a coefficient of variation $1/\sqrt{k}$; the genome-wide test uses $k = 2$, the smallest departure from memorylessness, and the profile analysis extends k to 32. For allele-resolved data the count is the sum over the two alleles, computed as the exact convolution of the single-allele laws; for the X chromosome the one-copy law is the single-allele law placed on the two-allele count grid.

Exact transient likelihoods For a promoter observed over a labeling window of length T, the probability of n new molecules is obtained by solving the chemical master equation of the promoter-plus-RNA system from the stationary distribution of the promoter at the start of the window, with RNA counted only if synthesized inside the window. We compute this distribution exactly, to floating-point precision, by uniformization of the generator on a truncated state space whose truncation error is bounded; the construction and its error bounds are given in the companion paper [28]. The likelihood of a gene is the product over cells of the probability of its observed new count given the cell's total-count depth. Maximum-likelihood fits used multiple starts and a convergence criterion on the gradient; a gene is evaluable when both models converged.

Exact p-values and the class calls The test statistic is twice the log-likelihood ratio of the multi-step model to the two-state model. Its null distribution at the fitted two-state parameters was computed exactly for 16,844 genes by a contour integral over the characteristic function of the per-cell contributions, and bounded by a Chernoff bound for the 14 genes whose contour computation exceeded the resource limit. Genes were called significant at a Benjamini-Hochberg false-discovery rate of 5 percent [35]. Three constructions then asked whether the statistic could arise without a third state. The two-state twin

of each gene simulates counts from the fitted two-state model at the gene's own depth with lognormal per-cell capture dispersion (coefficient of variation 0.3) and refits both models, giving the statistic a memoryless promoter of that gene's parameters produces under capture over-dispersion; the copy-number twin repeats this with the allele and copy-number structure of the data; the cell-cycle refit repeats the fit on the G1 cells alone, with cells assigned a phase by module scoring of S and G2/M gene sets, and requires the G1 statistic to exceed the twins' statistics on the same cells. A gene enters the refractory class when its statistic exceeds 8.6 and three times the larger of 4.3 and each twin's statistic, and survives the G1 refit. A gene enters the two-state class when its statistic is below -4.3, that is, when the two-state model is favored by the same margin. All other evaluable genes are undetermined. To assess the sensitivity of the class sizes to these criteria, we recomputed both the refractory and the clustered class with the margin over the two twin statistics set to 2x or 4x instead of 3x, with any two of the three controls (misspecification twin, copy-number twin, G1-only refit) required instead of all three, and with both relaxations together. The likelihood-ratio bar of 8.6, the 5 percent FDR requirement of the exact test, the treatment of a non-converged twin as a failed control, and the restriction of the clustered candidates to the 1,060 two-state genes were kept (Figure S1; Table S5 lists the membership of every gene under every variant).

Depth profiles and the clustered alternative For twenty genes with small count ranges, twelve from the refractory class and eight from the two-state class, taken at evenly spaced ranks of the likelihood-ratio statistic within each class, the likelihood was profiled over chain depth $k \in \{1, 1.5, 2, \ldots, 8\}$ with fractional depths defined through the Erlang shape parameter (refractory genes), or over the squared coefficient of variation of a two-timescale silent period on $[1, 2.9]$ (two-state genes); the depth profile was extended to $k = 16$ and 32 for eleven of the twelve refractory genes (the twelfth stopped on a wall-clock limit). The same twenty genes were also fitted with chains of depth 3 and 5 and with the clustered law; the resulting likelihood ratios are single-gene comparisons and are reported as log-likelihood gains, not as significance calls. The two-state genes were not profiled over chain depth; their chain-depth curves in Figure 1D are the fits at depths 1, 2, 3 and 5 from this twenty-gene comparison, evaluated with the observation model of the refractory profiles, which integrates over per-cell capture dispersion. Both classes are drawn in that panel as the difference in $2 \log L$ from the two-state model. A profile is flat above a depth when the change in $2 \log L$ across the next doubling is below the noise floor of the fit. The clustered alternative is a silent period with two exponential timescales (a hyperexponential mixture), fitted with the same active state. Chain depth is not rankable from a single window: with the chain extended beyond the tested range the likelihood is flat for ten of eleven refractory genes, so depths are reported as lower bounds. Twelve further clustered genes were added to the panel by a rule fixed before the fits: the clustered genes meeting the same count-range rule (34 of the 329, one already in the panel), at twelve evenly spaced ranks of the clustered statistic, each fitted with the same five models under the same observation model; the comparison with the seven telegraph genes is the one-sided Mann-Whitney test on the depth-5 differences, stated in advance with its 0.05 bar.

Class summaries and expression matching Per-gene mean silent and active periods are $1/k_{\mathrm{on}}$ and $1/k_{\mathrm{off}}$, burst size is $k_{\mathrm{syn}}/k_{\mathrm{off}}$, duty cycle is $k_{\mathrm{on}}/(k_{\mathrm{on}} + k_{\mathrm{off}})$, mean expression is $k_{\mathrm{syn}}/k_{\mathrm{d}}$ times the duty cycle, and the half-life ratio is the ratio of the silent period's half-life, $\ln 2/k_{\mathrm{on}}$, to the mRNA half-life, $\ln 2/k_{\mathrm{d}}$. Class contrasts are Mann-Whitney tests with Cliff's delta as effect size. Expression-matched pairs were formed one-to-one without replacement within a caliper of 0.1 dex in mean expression (156 pairs for the allele analysis; 186 pairs for the motif, Lisa, and ChIP analyses at a caliper of 0.15 dex). The fraction of refractory genes silent for longer than a cell cycle is given at two thresholds, 24 hours (165 of 186 genes) and a 20-hour fibroblast cycle (171 of 186).

Gene Ontology enrichment Enrichment of the 186 refractory genes was computed with GOATOOLS (propagated counts, Benjamini-Hochberg correction) against a background of the 12,243 evaluable genes with a GO annotation, using the MGI annotation file and the go-basic ontology. In the primary analysis, genes annotated to cell cycle (GO:0007049) or any of its is-a or part-of descendants were removed from the study set (164 of 186 refractory, 299 of 329 clustered, and 620 of 731 telegraph genes remained); the secondary analysis retained them (171 annotated refractory genes); the Results quote the primary analysis, and the secondary gives the same two terms at 4.6- and 7.6-fold ($p = 1.3\times10^{-22}$ and 2.0×10^{-12}). The clustered class and the 731 telegraph genes were analyzed with the same settings and background, and contrasts between classes on a given term are two-sided Fisher's exact tests on the same propagated annotations. A gene counts as a secreted or matrix protein when it is annotated to extracellular region (GO:0005576) or extracellular matrix (GO:0031012), or to an is-a descendant of either, in the same propagated annotation; percentages are of genes with any Gene Ontology annotation (171 of the 186 refractory genes). Similar expression means a fitted mean expression within the 5th to 95th percentile range of the refractory class, 0.42 to 10 transcripts per cell; the 1:1 expression-matched partners formed by the rule above give the same result (13 of 186 two-state partners and 9 of 186 undetermined partners; Fisher's exact test $p = 2 \times 10^{-13}$ and 3×10^{-14}). Promoter methylation and CpG-island contrasts between classes were tested on expression-matched pairs formed by the rule above (Mann-Whitney on measured promoters and Wilcoxon signed-rank on pairs with both members measured for methylation; Fisher's exact test and McNemar's test for CpG-island overlap).

Promoter chromatin features Per-gene features were computed on GENCODE vM25 (mm10): promoter GC content and CpG observed-to-expected ratio over TSS −300 to +100 bp, CpG-island overlap, mean signal of five histone marks over TSS ±1 kb, and mean CpG methylation from MEF whole-genome bisulfite sequencing over TSS ±500 bp, undefined for promoters with no covered CpG. Spearman correlations between the fitted silent and active periods and each feature were computed within each class, with partial correlations from rank-regression residuals controlling for mean expression. Class-specific regressions of each fitted rate on the standardized feature set (refractory, telegraph and clustered genes, each on the rates of its own model) used heteroskedasticity-consistent standard errors; class differences in a coefficient were tested in the pooled model with class-by-feature interactions (telegraph as reference, a two-degree-of-freedom Wald test per feature and rate), Bonferroni-corrected over the 21 feature-rate pairs with methylation (genes with measured methylation) and the 18 without it (all classed genes).

Motif scan Promoter sequences (TSS −500 to +100 bp) of the 186 refractory genes and their 186 expression-matched two-state genes were scanned with 1,019 JASPAR CORE vertebrate position frequency matrices (the latest version per base matrix identifier) using log-odds scores. Each promoter's hit threshold for each motif was set at the 99th percentile of the best scores over dinucleotide-preserving shuffles of the promoter's own sequence, pooled across the class (930 shuffle draws), so that base composition is controlled by construction. The class contrast per motif is the difference in hit fraction, with significance by 10,000 label permutations within matched pairs and Benjamini-Hochberg correction. A surviving motif is labeled composition-driven when its consensus is GC-rich (GC ≥ 0.7) or contains CpG.

Lisa The two gene lists were submitted to the Lisa web service[38] with default settings; the ranked factor lists were read for the top-ranked ChIP-seq factors on each side.

ChIP-seq metagenes and the fixed predictions Public ChIP-seq tracks for MED1, MYC, and RNA polymerase II in mouse embryonic fibroblasts and for BRD4 in NIH/3T3 cells were downloaded as bigWig files (accessions in Table S1). Experiments with fewer than a minimum peak count or a median occupancy below a floor were excluded before any contrast was computed. Signal was binned at 25 bp from TSS −1 kb to +3 kb in the direction of transcription. Occupancy is the mean signal over TSS

±500 bp; the pausing index is the mean signal over TSS −50 to +300 bp divided by the mean over +1 to +3 kb, undefined where the gene-body signal is zero. Five predictions were written down with their direction before the tracks were examined: that BRD4, MYC, and polymerase occupancy would be lower at refractory promoters, that MED1 would be higher, and that the pausing index would be higher. Each contrast is the pooled within-pair difference across usable experiments, with significance by permutation of pair labels and Benjamini-Hochberg correction across the five tests; sign consistency is the fraction of experiments agreeing with the pooled sign. The duty-cycle comparison and the factor-to-polymerase ratio were computed after the results were known and are labeled post hoc in the text; the pausing-index result was recomputed on the 142 pairs with the index defined on both sides.

Co-bursting and inter-allele correlation A gene bursts in a cell when its new count is positive. For every pair of classified genes whose mean silent periods exceed twice their mRNA half-lives, the co-bursting statistic is the Mantel–Haenszel log odds ratio between the two genes' bursting indicators across cells, stratified by decile of the cell's total molecule count (Haldane's 0.5 added to the pooled numerator and denominator); the class-level contrast is the median over refractory–refractory pairs minus the median over refractory–two-state pairs, compared with a null from permuting each gene's indicator across cells within each depth decile, which preserves each gene's bursting frequency and depth profile and removes any dependence between genes. For the allele analysis, the inter-allele overdispersion of a gene is the beta-binomial intra-class correlation of the within-cell split of new molecules between the two alleles, given the cell's total, fitted by maximum likelihood over cells with at least one allele-resolved new molecule. A higher value means a more extreme split than binomial and so less correlation between alleles. Genes were compared across classes on 156 expression-matched pairs, on the unmatched sets, and with partial tests controlling for allelic imbalance, molecule count, and degradation rate. Because the allele-resolved unit count per new molecule differs between classes (medians 4.22 in the refractory class and 1.67 in the two-state class), the absolute values are not interpretable as correlation levels; the class contrast is the reported result.

Labelled reads along the gene To compare nascent RNA between gene groups within the same cells, the first 2×10^8 read-1 records of the NASC-seq2 run were realigned to mm10 with bwa mem (adapter-trimmed, mapping quality ≥10; 98.6 percent mapped) and reads overlapping the span of 1,245 target genes were kept (the two archetype neighbourhoods in both membership definitions, the refractory genes, their matched partners and 223 further two-state genes); positions shared with any other annotated gene were masked because internal reads are unstranded. A read was called new when its alignment carried at least one T-to-C or A-to-G mismatch, assigned to exon or intron by the GENCODE vM25 exon union of its gene, and to a decile of the gene body by its midpoint. Per gene we report the intronic fraction of new reads (raw and normalised by intron and exon length), the fraction of new reads, and the slope of log new-read density over deciles minus the same slope for all reads, a within-gene measure of how far $5'$ the new RNA sits relative to the old; groups were compared as for the other features.

Nascent transcription (PRO-seq) Strand-specific, normalized PRO-seq coverage of untreated immortalized wild-type mouse embryonic fibroblasts was taken from two studies: GEO GSM1843438 (no heat shock)[46] and GSM3665947 (unconditioned, spike-in normalized, replicates combined)[47], both mm10. For every gene of at least 3.5 kb we took the mean absolute sense-strand signal in the promoter-proximal window (TSS to +300 bp), in the gene body (+1 kb to 0.5 kb before the annotated end, capped at 100 kb), in the first and last quarters of that body window, and the antisense signal in the 1 kb upstream of the TSS; the pausing index is the promoter-proximal over the body density. Refractory genes were compared with their expression-matched two-state partners by the Wilcoxon signed-rank test and by Cliff's delta, and the two two-state archetype neighbourhoods with each other in the same way, in each dataset separately.

Allele coupling by bivariate mixing Given the promoter path, an allele's new-RNA count in the window is Poisson with mean $\beta_c\theta$, where $\theta = \int_0^W r\, 1_{\mathrm{on}}(t)\, e^{-k_d(W-t)}\, dt$ is the capture-free window activity and β_c the cell's capture, so each allele is a mixed Poisson whose mixing law is fixed by the fitted kinetics. Following a construction of M.H., the joint mixing density of the two alleles was taken as $f(\theta_1)f(\theta_2)[1+\omega(\theta_1-\mu)(\theta_2-\mu)]$, for which $\mathrm{corr}(\theta_1,\theta_2) = \omega\sigma_\theta^2$ and, because $\int \beta\theta\, \mathrm{Poi}(n \mid \beta\theta) f(\theta)\, d\theta = (n+1)P(n+1)$ for any mixed Poisson, the joint count law is $P(n_1)P(n_2)[1+\omega g(n_1)g(n_2)]$ with $g(n) = (n+1)P(n+1)/(\beta P(n)) - \mu$, P being the allele's marginal count law in the cell's depth bin. Per gene, f was the production per-copy law (multi-step for refractory, two-state otherwise) at the fitted rates, marginalised over old RNA; μ and σ_θ^2 are its first two factorial moments at unit capture. The deposited allele tables count new-RNA UMIs (a median 3.5 units per reconstructed molecule), so each allele carries a unit yield κ estimated from its mean. Three estimates of the coupling were computed over the 8,912 allele-called cells: the moment estimate $\hat\rho = \sum_c (n_{1c} - s_c\kappa_1\mu)(n_{2c} - s_c\kappa_2\mu)/(\kappa_1\kappa_2\sigma_\theta^2 \sum_c s_c^2)$ with s_c the cell's relative capture, its variant with σ_θ^2 replaced by the observed excess variance of depth-normalised units, and the maximum-likelihood ω of the joint count law over its admissible interval (density non-negative on the support); significance per gene is from 200 permutations of one allele's counts across cells within depth bins. The same quantities for the reference allele of each refractory gene against the reference allele of its expression-matched two-state partner, in the same cells, give the cell-level floor, and X-linked genes serve as a negative control. The same moment estimate was computed for all classed autosomal genes, and cell-cycle-periodic genes (the S and G2/M marker set used for cell-cycle scoring) were compared with the rest by a one-sided Mann–Whitney test, also after regressing the estimate on the logarithm of the gene's allele-unit count.

Coupled-allele law To test whether inter-allele coupling can be fitted directly, we built an exact transient law for two alleles sharing a cell-level gate: the gate opens at rate g_{on} and closes at rate g_{off}, each allele switches on at k_{on} only while the gate is open and off at k_{off}, and synthesis and degradation are as in the two-state model. The joint law of the two alleles' new counts after the labeling window is computed by uniformization on the state (gate, allele states, two counts) from the stationary promoter start, thinned per allele by the cell's capture probability, and the independent-allele model is the nested case $g_{\mathrm{off}} = 0$. The law reduces to the product of single-allele laws in that limit to 5×10^{-15}, and an event-driven simulator independent of it reproduces the joint law to within Monte Carlo error. Planted truth at the class-median operating points (8,916 cells, real capture bins, 20 replicates per setting, gate rates 0.01 to 0.05 per hour) gave a likelihood-ratio detection rate of 100 percent at the two-state point and 25 to 55 percent at the refractory point, with no false detections without a gate; under a real gate the independent-allele fit under-estimates the burst-on rate about two-fold, a model dependence we note for the fitted rates. The real-data fit was not run under the analysis plan, which required recoverability at the refractory point. The decomposition of two-copy covariance into allele-intrinsic and allele-extrinsic parts follows the classical intrinsic and extrinsic split [95] in the form used for nascent-RNA counts of two alleles [112]. The two-copy coupling likelihood of Zhang, Bao and Xu multiplies the activation rate of a two-state copy [89], so the coupling in this law is not comparable to it.

Noise and information rate Stationary first and second moments of mRNA and protein for the two-state and Erlang-k promoters were computed exactly from the moment equations of the promoter-mRNA-protein system with a protein half-life of ten hours and validated against the numerically solved chemical master equation to a relative error of 2×10^{-13}. The information rate of the burst train is the entropy rate of the promoter-state point process at dwell-matched parameters.

Promoter and enhancer feature regression Per gene, log10 burst-on rate, active-exit rate, synthesis rate, burst size (synthesis over active-exit), and duty cycle from the fitted rates were regressed on standardized feature blocks by ridge regression with the penalty chosen by inner cross-validation, in five-

fold cross-validation repeated twenty times; a gradient-boosted tree model (200 iterations) was run beside it as a nonlinear check with five repeats. The promoter block is GC content, CpG observed-to-expected ratio, CpG-island overlap, and TATA presence in the −300 to +100 window, H3K4me3, H3K27ac, H3K27me3, H3K9me3, and H3K36me3 within 2 kb of the transcription start site, and whole-genome bisulfite methylation of the promoter (median-imputed with a missingness indicator where unmeasured); the enhancer block is the number of enhancers within 100 and 500 kb, super-enhancer association and distance to the nearest super-enhancer (genes with none within 500 kb take the window edge), and Hi-C compartment eigenvector, insulation, and contact density; log gene length, exon count, and transcript count enter every model. The gene set is the 16,214 autosomal genes with complete features among the 16,858 with converged fits. Out-of-fold R^2 is summarized over repeats, whose spread reflects fold assignment only; the positive control was the model-implied mean expression, predicted at R^2 0.41. A version with expression as a covariate was run and discarded because the model-implied expression is a function of the outcomes.

Enhancer activity and the between-cell-type comparison Enhancer activity per gene is the H3K27ac ChIP-seq signal of mouse embryonic fibroblasts (Chronis et al.[113], ChIP-Atlas uniform processing of experiment SRX2396384, the track behind the promoter H3K27ac feature) over ENCODE SCREEN distal enhancer-like elements (registry V3, mm10, 293,072 elements) whose midpoints lie within 100 kb of the canonical transcription start site and more than 2 kb from it (5 kb as a sensitivity run, which moved no correlation by more than 0.01): per element the mean coverage, per gene the sum of mean times length over linked elements, entered as $\log_{10}(1+\text{sum})$ together with the log maximum element signal and the log number of linked elements above the 90th percentile of all elements. These three features form the activity block added to the regressions above (same genes, folds and repeats). For the promoter-features-at-enhancers comparison the whole promoter feature set was re-measured over the same linked elements: GC fraction and CpG observed-to-expected ratio of the element sequences (plain and H3K27ac-weighted means), the fraction of linked elements overlapping a CpG island, the five MEF histone marks (per element mean coverage; per gene $\log_{10}$ of one plus the length-weighted sum and of one plus the maximum), MEF methylation (mean percent over covered cytosines in the linked elements, 112,411 of 293,072 elements covered, median-imputed with an indicator for the 23 percent of genes without coverage), the log number of linked elements, and the 204 ChIP-Atlas MEF antigens of the promoter regression re-scored at the elements (per experiment the strongest peak overlapping each element, per gene $\log_{10}$ of one plus the maximum over linked elements, per antigen the mean over experiments; 3,378 of 3,394 peak files available; 199 non-constant), 219 features in all; ridge as above with twenty repeats for the headline blocks and five for the descriptive ones, a boosted-tree check, and a Lasso ranking; enhancer-RNA features come from a realignment (bwa mem, MAPQ ≥ 10) of the first 0.9×10^9 raw NASC-seq2 read-1 records, fetched afresh from the European Nucleotide Archive after the local copy proved corrupt, keeping reads on all 293,072 distal elements; per gene, reads per kilobase over the linked elements that do not overlap any GENCODE gene span extended by 2 kb (71,130 intergenic elements), an any-read indicator, the fraction of those reads carrying a T-to-C or A-to-G mismatch, and the element counts; rank correlations carry bootstrap intervals over genes (2,000 resamples) and the partial rank correlation given expression is computed on rank residuals. For the between-cell-type comparison the per-gene burst frequency and burst size of Larsson et al.'s fibroblasts and embryonic stem cells are the exact profile-likelihood re-fits of their CAST-allele counts (the embryonic stem cells on the 94 unique cells of the deposited table), 4,553 genes present in both with positive estimates; the change in enhancer activity is the difference of the $\log_{10}(1+\text{sum})$ feature between the fibroblast track and the embryonic-stem-cell track of the same study (SRX2396387), which are coverage-normalized by one procedure so that a global scale enters only as an offset. Reported are Spearman correlations and least-squares R^2 with bootstrap intervals, the joint standardized regression of the activity change on both kinetic changes, the contrast of the 100 largest gains against the 100 largest losses, the subset

whose corrected 95 percent intervals do not overlap between cell types, and the rolling median over 200 genes of the activity change along each kinetic ordering, summarized by its range in units of the per-gene standard deviation. The pairing is by cell type rather than by cell line (embryonic fibroblasts and an embryonic-stem-cell line for the ChIP-seq; adult tail fibroblasts and hybrid embryonic stem cells for the kinetics), linking is by distance rather than by measured contact, and the enhancer set and linking rule of the original figure are not restated here, so the comparison reproduces its design rather than its elements.

Chromosome-resolved co-transcription of neighbouring genes For each autosomal gene with at least 50 cells positive on each allele (12,557 genes), the new-RNA indicator (labelled count above zero) was formed separately for the reference and the alternative chromosome across all 8,916 cells. For every same-chromosome pair the Pearson correlations of the indicators were computed for the two same-copy combinations and the two opposite-copy combinations, from per-chromosome products of the standardized indicator matrices, and the statistic is the mean same-copy minus the mean opposite-copy correlation; any cell-level covariate enters the two identically. Its zero was checked on 20,000 random different-chromosome pairs. Pairs were binned by the distance between transcription start sites; pairs with overlapping gene bodies were excluded and head-to-head pairs within 10 kb reported separately as a check against shared transcripts. Intervals are bootstrap percentiles over pairs (2,000 resamples; a gene-resampling bootstrap, which respects the sharing of genes among pairs, gave intervals about twice as wide and the same conclusions) or the normal approximation above 50,000 pairs. Whole-chromosome allelic imbalance was screened per cell and chromosome: with at least 20 allele-assigned new-RNA reads on the chromosome, a reference fraction below 0.2 or above 0.8 flags the cell for that chromosome (595 cells flagged on at least one chromosome; 16 to 92 cells per chromosome), and the reported numbers mask flagged cells chromosome by chromosome; without the mask a distance-independent same-copy excess of 0.0014 is present at all separations and falls to 0.0004 with it, while the short-range component above that level is unchanged (0.0028 before, 0.0026 after, within 10 kb). Restricting to the 6,290 G1 cells changes no conclusion. Contact domains are the intervals between insulation boundaries called by cooltools (25 kb bins, 100 and 250 kb windows, default threshold) on Hi-C of L929 mouse fibroblasts (GEO GSM8517974), the map used for the Hi-C features above; same-domain and different-domain pairs were compared within distance bins. Kinetic similarity of neighbours is the symmetrized Spearman correlation, across pairs in a distance bin, of the two genes' fitted log rates from the parameter set above, and refractory concordance is the odds ratio, with Haldane's correction, of both-refractory pairs among classed pairs in a bin against classed pairs more than 5 Mb apart. The imprinted pairs named in the Results were not supplied to the analysis; they were recognised among the 25 pairs with the largest same-copy excess by their membership in the maternally expressed *Dlk1–Dio3* domain[59] and the paternally expressed Prader–Willi domain[60].

ChIP-Atlas feature blocks All 3,394 MEF experiments in the ChIP-Atlas mm10 experiment list of 2026-09-08 in the classes TFs and others (163 antigens), Histone (38), RNA polymerase, ATAC-seq and DNase-seq were used; the other fibroblast lines (1,307 experiments, 122 antigens) form a separate block. A gene's value is $\log_{10}(1 + \text{score})$ of the strongest MACS2 peak ($q < 10^{-5}$) within 1 kb of its transcription start site, averaged over an antigen's experiments; 21 of the 4,701 peak files were empty and were skipped. Blocks of 16 (the promoter and gene-structure features of the previous paragraph), 54 (plus histone marks), 217 (plus factors) and 220 (all) columns were fitted as in the previous paragraph. A Lasso with its penalty from inner five-fold cross-validation was fitted in the same 100 fold-fits and once on all genes. A covariate register refitted the base and full blocks with log mean expression appended, same folds and seeds, to separate chromatin information from expression. Refractory and partner promoters were compared per antigen (Mann-Whitney, Benjamini-Hochberg) on the 124 pairs whose partner is a telegraph gene and on all 186 pairs. For each antigen the score is averaged over studies,

genotypes and treatments. The repressed comparison set is drawn from the 5,702 protein-coding genes on chromosomes 1–19 (GENCODE vM25) that were never detected in the 8,916 cells: each refractory gene was paired, each repressed gene used once, with the repressed gene nearest in promoter CpG observed-to-expected ratio (within 0.05; 175 autosomal pairs for the histone marks and methylation, and, requiring the same CpG-island status as well, 186 pairs for the enzyme occupancies and Figure 4), and refractory and repressed promoters were compared per feature by the Mann-Whitney test. Enzyme occupancy (Figure 4B) is the same peak score taken from the ChIP-Atlas target-gene table of each protein (peaks within 1 kb of the TSS, all mouse cell types pooled, zero for a gene absent from the table), compared between refractory promoters and their expression-matched two-state partners by the Wilcoxon signed-rank test.

Single-cell chromatin accessibility Fragment files were taken from GEO: 10x scATAC-seq of untreated primary 129/Ola embryonic fibroblasts (GSE160764, sample GSM5852360, 0 h; 1.78×10^8 fragments)[49] and 10x Multiome (RNA and ATAC in the same fixed cell) of an immortalized Cas9-expressing embryonic-fibroblast line carrying a CROP-seq guide library (GSE297418, replicates GSM8990892 and GSM8990896, with the series' raw RNA count matrix and cell metadata)[50]. For the scATAC sample, barcodes with at least 1,000 fragments were kept as cells (12,124; median 11,588 fragments; a threshold of 5,000, leaving 9,149 cells, changes no result); for the Multiome series only the 1,935 cells whose guide targets nothing or a gene not expressed in fibroblasts (non-targeting, *Rgr*, *Opn4*, *Neurod1*, *Olig1*, *Rrh*, *Eomes*) were used, matched to fragments by barcode and replicate (median 15,550 fragments and 17,250 RNA counts per cell). For every classed gene on chromosomes 1–19 (175 refractory, 1,024 two-state) and 1,000 random silent protein-coding genes, a fragment was assigned to a promoter when its midpoint fell within 1 kb of the transcription start site, and the promoter was called accessible in a cell with at least one such fragment. Per gene we recorded the accessible fraction of cells (also averaged over cell-depth deciles), the mean fragment count, and a zero-inflated Poisson fit $n_{gc} \sim \pi_g\,\delta_0 + (1-\pi_g)\,\mathrm{Poisson}(\lambda_g d_c/\bar{d})$ with the cell's total fragment count d_c as offset, maximized over (π_g, λ_g) by Nelder–Mead, together with twice the log-likelihood gain over the Poisson fit. Refractory and two-state genes were compared within the expression-matched pairs of the main analysis by the Wilcoxon signed-rank test, within pairs rematched on the accompanying scRNA-seq (GSM5852363) or on the Multiome RNA (nearest two-state gene within 0.1 dex of mean count, each used once), and within pairs matched on mean promoter accessibility (0.05 dex); the zero-inflation likelihood ratio was also compared across all classed genes after regressing its rank on the rank of mean accessibility (Mann–Whitney on residuals). Rank correlations of the accessible fraction with the fitted rates are Spearman, and partial values are correlations of rank residuals after regression on ranked log expression. In the Multiome control cells each gene–cell pair was classified by transcript present or absent and promoter accessible or not; per gene we took the fraction of cells with neither, the fraction without an accessible promoter among cells without transcript, and the Mantel–Haenszel odds ratio of transcript presence on accessibility over terciles of cell fragment depth (0.5 added to each cell of the table).

Clustered fit of the two-state class The clustered promoter has a silent period drawn from a two-component hyperexponential mixture with exit rates $(2 \pm \sqrt{2})\,a$ and equal weights, so that its mean silence equals that of a two-state promoter with burst-on rate a and its squared coefficient of variation is fixed at 2, below the 2.3 to 4.3 measured for a clustered promoter in live cells[26], and the same active state; it adds no free parameter beyond the two-state model, and the two are compared by twice the log-likelihood ratio on the same exact transient likelihood. Every gene with a converged two-state fit was refit under the clustered law on the same count grid (genes whose grid exceeded the cap used for the twins, 135, are listed as not fitted), but the clustered class is called within the two-state class only: a gene whose two-state description is itself undetermined cannot be said to prefer a clustered silence over a memoryless one. Of the 1,060 two-state genes, 23 were not fitted under the clustered law and are counted

as not clustered. The 731 genes of the two-state class not called clustered form the telegraph class. Calls were hardened by the rule used for the refractory class: statistic above 8.6; above three times the larger of 4.3 and the misspecification twin (a two-state gene at the fitted rates refit under the clustered law) and of the copy-number twin (a two-state gene with the cell-cycle copy-number distribution); and survival of a refit on G1 cells. Because the two-state and refractory classes are disjoint, no gene can satisfy both rules. Of the 329 clustered genes, 287 were called on the converged clustered statistic and 42 on the per-gene fallback statistic used where the clustered fit did not converge. The same rule applied to the 15,612 undetermined genes would pass 37 of them; these are not reported as a class. A 964-gene pilot preceded the full run and is superseded by it. Weight robustness: the 1,033 fit-able two-state genes were refit under the clustered law with the fast-component weight at 0.4 and at 0.75 (exit rates set so that the mean silence and its squared coefficient of variation of 2 are unchanged; below a weight of 1/3 no such rates exist), by the same protocol and on the same grids as the equal-weight fit, and the class rule was re-applied at each weight with the equal-weight twin nulls and G1 outcomes; 27 genes above the grid cap were not refit. Keep, enter and likelihood-profile counts are reported in the next paragraph; genes whose refit did not converge at a new weight count as lost. Shape robustness: the same 1,033 genes were refit at the fast weight 0.75 with the silence's squared coefficient of variation set to 3 and to 4 (exit rates a/x and a/y with $x = 1 - \sqrt{(1-w)(c-1)/(2w)}$ and $y = 1 + \sqrt{w(c-1)/(2(1-w))}$ for shape c, admissible while $c < (1+w)/(1-w)$), by the same protocol; the class rule was re-applied with the shape-2 twin nulls, so the counts of clustered genes that stay are comparable across shapes while the counts of telegraph genes that would enter are upper bounds until the twin controls are refit at each shape.

Robustness of the clustered class to the weight and shape conventions The assumption of equal probability for the two states is part of our model, and it is not a parameter fit by the data, and as a result the slow component holds 85 percent of the silent time (Figure 2F). The membership of the class depends on that assumption: refitting every two-state gene with the fast-component weight moved to 0.4 or 0.75, with the mean and dispersion of the silence held, keeps 156 and 260 of the 329 clustered genes (118 at all three weights) and admits 9 and 202 of the 731 telegraph genes, while the likelihood itself is highest at 0.75 for 157 of the 329, at the equal weight for 134 and at 0.4 for 38. The clustered class is therefore the class called under the equal-weight law, and the equal weight is a convention rather than a value the counts single out: they lean toward a heavier fast component, under which the class would be larger. The same holds for the shape: refitting every two-state gene at the fast weight 0.75 with the squared coefficient of variation of the silence raised from 2 to 3 and to 4 (the range measured for a clustered promoter in live cells) keeps 326 and 321 of the 329 clustered genes, and the likelihood of a clustered gene is highest at 4 for 169 of the 329 and at 3 for 154, at the fixed value 2 for only 6 (median gain over the fixed shape 12 nats at 3 and 15 nats at 4), so the fixed shape is the conservative end of what the counts allow rather than their preference; how many telegraph genes would join the class at those shapes needs the control refits at the same shape and is not yet stated.

Change detection Bursts were modeled as a renewal process with interburst intervals exponential (two-state) or Erlang-k at the same mean. A step change reduces the mean interval from T_0 to T_1 at a random time. The Shiryaev-Roberts statistic $R_n = (1 + R_{n-1})\Lambda_n$, with Λ_n the likelihood ratio of the nth interval under the post- and pre-change laws, is the limiting form of Shiryaev's Bayesian detector[68] for a vanishing change prior[69]. Its threshold was set by bisection on pre-change-only runs to a mean time to false alarm of 10, 100, or 1,000 times T_0. The delay is the mean time from the change to detection over realizations in which detection occurred after the change, and the fraction of such realizations is reported. The operating points are TFF1's measured interburst times at the half-maximal dose and at saturation[23]; the wide separation is a 24-fold step at the same interval scale. A Shiryaev detector with the change prior included was run alongside and agrees with the Shiryaev-Roberts delays to within

0.4 to 8 percent. The data-level check used the published dose-response statistics for TFF1, modeled as a two-timescale mixture, against a mean-matched Poisson process at twelve settings of deep-state weight and false-alarm budget. The refractory-against-clustered comparison used the same detector on a hyperexponential clustered law (slow-component weight 0.05, 0.15, or 0.30; slow-to-fast mean ratio 10 or 180; fast mean 185 minutes) and on Erlang-2, -5, and -7 and exponential laws at the identical pre- and post-change means, for a 2.8-fold and a 24-fold change, at mean times to false alarm of 100 and 1,000 pre-change intervals, under two change conventions fixed in advance: every component's mean scaled by the same factor, or the fast component alone changed with the slow component fixed; 40,000 realizations per run, and the adjudication on Erlang-5 against the clustered law at every setting. Under the first convention (every component rescaled, 185 to 66 minutes), the clustered law of the main text (squared coefficient of variation 2) is detected 1.3 to 1.5 times later than the exponential law and 4.8 to 6.8 times later than Erlang-7 at mean times to false alarm of 100 and 1,000 intervals (20,000 realizations), and its Kullback–Leibler divergence per interval is 0.62 of the exponential value.

Change detection on counts In this setting the observable is the mRNA count sampled every 0.1 mean lifetimes, the promoter state is hidden, and the joint chain of promoter state and count (two-state with burst-on and active-exit rates of 0.1 and 0.9 or 1 and 1 per lifetime; three-step refractory at the same duty cycle) is propagated by the exact frame kernel $\exp(Q\Delta)$ on a truncated count grid of up to 400 transcripts. The per-frame likelihood ratio between the post- and pre-change laws is the ratio of the two exact filters' predictive probabilities of the observed count, and the Shiryaev-Roberts threshold was set by bisection to a mean time to false alarm of 100 lifetimes. The change occurs at a geometric time with mean 10 lifetimes and multiplies the output mean by 1.5 or 3 through one parameter at a time (synthesis rate, burst-on rate, or active-exit rate). The burst-on and active-exit channels are undefined where the post-change duty cycle would exceed 0.95 (fast switching at 3-fold), which leaves 40 runs of 10,000 realizations. Delays are conditional on detection after the change (fraction 0.92 to 0.996; false alarms before the change 0.4 to 8 percent). The comparison fixed in advance was the three-step promoter against the two-state one on the burst-on channel, the fastest channel for the two-state promoter. The per-snapshot Fisher information of the stationary count law was computed for the two-state promoter at the same operating points by central differences in log-parameters, the logarithmic parametrization of Fox and Munsky[114]. In that form it is invariant to the parametrization; in raw rate units the same quantity carries a factor $1/\theta^2$ that inverts the ordering.

Noise and detection time in closed form For a promoter with one active state left at rate k_{off}, a silent period T_{off} of mean $1/k_{\mathrm{on}}$ with Laplace transform $\hat{f}(\lambda) = \langle e^{-\lambda T_{\mathrm{off}}} \rangle$, synthesis rate k_{syn} and decay rate k_{d}, the stationary Fano factor is $F = 1 + k_{\mathrm{syn}}/(k_{\mathrm{d}} + k_{\mathrm{off}}\varphi) - k_{\mathrm{syn}} k_{\mathrm{on}}/(k_{\mathrm{d}}(k_{\mathrm{on}} + k_{\mathrm{off}}))$ with $\varphi = 1 - \hat{f}(k_{\mathrm{d}})$, where $\hat{f}(k_{\mathrm{d}}) = (k\,k_{\mathrm{on}}/(k\,k_{\mathrm{on}} + k_{\mathrm{d}}))^k$ for a k-step silence and $2k_{\mathrm{on}}(k_{\mathrm{on}} + k_{\mathrm{d}})/(2k_{\mathrm{on}}^2 + 4k_{\mathrm{on}}k_{\mathrm{d}} + k_{\mathrm{d}}^2)$ for the clustered law. The formula was derived twice, from the moment equations and from the variance decomposition of the doubly stochastic Poisson process, and agrees with the exact moment solver above to a relative error of 8×10^{-12} over the 1,246 fitted rate sets. With bursts resolved, the delay in detecting a step $k_{\mathrm{on}} \to k'_{\mathrm{on}}$ at mean time to false alarm γ is $\ln \gamma / I$ to first order[70], where I is the Kullback-Leibler rate of the post-change against the pre-change cycle: $I = k[\ln(k'_{\mathrm{on}}/k_{\mathrm{on}}) + k_{\mathrm{on}}/k'_{\mathrm{on}} - 1]/(1/k'_{\mathrm{on}} + 1/k_{\mathrm{off}})$ for a k-step silence, and a one-dimensional integral for the clustered law (0.59 of the memoryless value at a 1.5-fold step). Per-gene delays were evaluated at each classed gene's fitted rates for a 1.5-fold step at $\gamma = 100$ cycles, with two steps for the refractory class (the fitted model), and the first-order value was scaled by the factor a per-cycle Shiryaev-Roberts simulation gives at that budget (0.39 for one step, 0.51 for two, 0.58 for three, 0.31 for the clustered law); the rank correlations are the same with or without the factor. At the false-alarm budgets of the count-level calculations above (9 to 50 pre-change cycles) the first-order formula overstates the gain, and on the slowest-switching rows the delay is set by the false-alarm clock rather than by the information rate.

Environment encoding A two-state environment with equal switching rates $1/\tau$, $\tau \in \{1, 10, 100\}$ mean lifetimes, was coupled to each promoter as a hidden coordinate of one joint chain (environment, promoter state, count) propagated by the exact frame kernel $\exp(Q\Delta)$, $\Delta = 0.1$ lifetime, on a count grid of up to 400. In the high state one parameter is changed so that the stationary mean rises by $c \in \{1.5, 3\}$ from a low-state mean $m \in \{1, 10\}$: the synthesis rate, the burst-on rate, or the active-exit rate (the last two undefined where the high-state duty cycle would exceed 0.95). Two-state and three-step promoters were run at duty cycles 0.1 and 0.5 (burst-on plus active-exit rate equal to one per lifetime) and, descriptively, at the fitted class medians 0.052 (three-step) and 0.32 (two-state). The information rate is the mean over frames of $\log p(y_t \mid y_{<t}, e_{\leq t}) - \log p(y_t \mid y_{<t})$, both terms from exact filters on the joint chain, the first conditioned on the environment at the frame boundaries, in bits per lifetime, over 200 realizations of 2,000 frames after 200 frames of burn-in; its standard error is over realizations. Tracking error is the mean squared error of the marginal filter's posterior probability of the high state against the truth. The low-state Fano factor is exact from the stationary law of the low-state chain. The comparison of information rates used a two-standard-error criterion chosen after the fact; the tracking comparison is descriptive. The frontier fixed in advance also named detection delay against noise, which was not evaluated in this analysis.

Archetype analysis Per gene, the fitted burst-on, active-exit, and synthesis rates and the published degradation rate were log-transformed and standardized over the 1,246 classed genes. Principal convex hull analysis[79] was run at three to six archetypes with ten optimizer restarts, keeping the best; significance is the t-ratio of the polytope volume to its column-shuffled controls, which is defined only for five or fewer archetypes in four dimensions because six vertices span a five-simplex of zero volume there. Archetype stability was assessed under two hundred bootstrap resamples, and each candidate task (detection delay, information rate, protein noise) was evaluated exactly at every archetype and at two hundred genes spanning the cloud. A first pass of the analysis with a single optimizer restart had produced a polytope with a detection archetype at the predicted position; the confirming corner was present in one sub-optimal restart (the seventh best of ten) and in no other, so single-restart archetype fits are not reliable at adjudication margins. The promoters do not form a significant polytope at any archetype number where the volume test is defined; a softer comparison of variance explained separates the real cloud from its shuffles only at three archetypes (51 against 38 percent), which the Results report as descriptive structure, and with two archetypes the segment from a generic two-state point to the refractory corner explains 29 percent of the variance (17 for shuffles). The refractory archetype is among the most stable under the bootstrap resamples. Adding promoter CpG content and methylation as fifth and sixth coordinates (612 genes with measured methylation; five coordinates on all 1,246) leaves the volume test inside the shuffled range at every archetype number ($p \geq 0.16$) while variance explained exceeds the shuffles at every number, as any set of mutually correlated coordinates will (CpG content correlates with the log burst-on rate at $r = 0.48$, methylation at -0.40); the refractory-rich archetype persists and acquires the low-CpG, high-methylation coordinates.

X-chromosome refit The 46 confidently classed X-linked genes were refit under the one-copy law with the same starts, seeds, and convergence criteria as the genome-wide two-copy fits, with two-copy controls run through the same fitting code; Tmsb4x, whose count grid exceeded the feasible size, was excluded. Per-cell X haplotype share and per-gene monoallelic fraction were computed from the allele-resolved new counts. The criterion that all ten refractory calls hold under the one-copy law, the twin-based bar of 4.3, and the escapee control were fixed before the refit. The cost of misspecifying the allele count was measured by a planted control at the fitted rates of Vegfd and Hs6st2: counts simulated from one-copy and from two-copy truths (two-state and three-state, two replicates each) and fit under both laws; a one-copy truth loses 11 to 43 log-likelihood units under the two-copy law and a two-copy truth loses 123 to 187 under the one-copy law.

Enhancer transcription The first 205 million reads of the raw read-1 archive of the NASC-seq2 fibroblast data (a byte-bounded head sample, random over cells and over the genome and nonrandom over flowcell tiles) were trimmed of the 19-base tag where present, aligned to mm10 with bwa mem, and kept at MAPQ ≥ 10 within the gene bodies and assigned enhancer windows of the 372 pair genes. Enhancer windows are intergenic 1-kb windows (no gencode vM25 gene body ± 2 kb) at or above the 95th percentile of the genome-wide intergenic MEF H3K27ac distribution, assigned to a gene when their center lies within 100 kb of its TSS; the per-gene analysis uses windows whose nearest annotated TSS is the gene's own, and the pooled analysis uses all windows within 100 kb of any pair gene of a class. The new-RNA signature is the fraction of reads carrying at least one T-to-C (or A-to-G) mismatch after `samtools calmd`, against G-to-C and C-to-G as the control substitution. The class contrast is the log2 ratio of pooled enhancer read density to gene-body read density, refractory minus matched, with a null from 10,000 within-pair label swaps and a power check from a planted two-fold thinning of the matched class's enhancer reads over 100 seeds; the per-gene analysis had four evaluable pairs at 20 reads and did not decide. The window set, the two analyses, and the branch rules were fixed before alignment, with one pre-alignment amendment: the window set was widened from the 10,000 strongest windows to the 95th percentile after a coverage scan showed the smaller set left 292 of the 372 genes without a window.

Statistics

Class contrasts of per-gene quantities use two-sided Mann-Whitney tests with Cliff's delta as the effect size; correlations are Spearman rank correlations with partial versions from rank-regression residuals; enrichment and multiple contrasts are corrected by the Benjamini-Hochberg procedure and reported as q; permutation tests use label permutation within matched pairs where pairs exist. Sample sizes (n genes, pairs, or experiments) are given with each result in the text and figure legends. Per-gene p-values for the model comparison are exact (or Chernoff-bounded, 14 genes) rather than asymptotic.

Supplementary information

Table S1 lists the ChIP-seq experiments used, with accession, antigen, cell line, peak count, median occupancy, and usability. Table S2 lists the 1,246 classed genes (186 refractory, 329 clustered, 731 telegraph) with fitted rates, class, and exact p-value. Table S3 gives the corrected burst-on rates for the X-linked refractory genes. Table S4 lists, for every gene that enters either class under any variant of the control criteria, its membership under each variant (Figure S1).

Acknowledgments

To be written.

Author contributions

Conceptualization, M.D.S. and M.H.; methodology and software, M.D.S.; formal analysis, M.D.S.; hypotheses on promoter function and archetypes, chromatin and BRD4, and the Lisa analysis, M.H.; writing, M.D.S. and M.H.

Competing interests

The authors declare no competing interests.

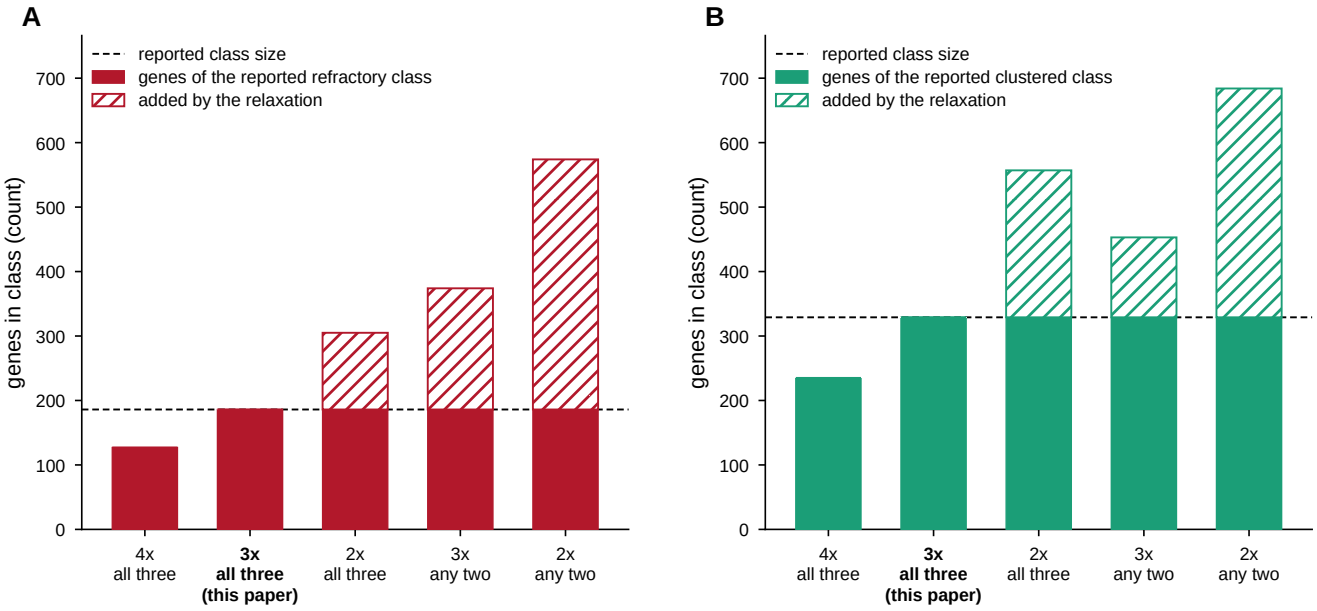

Figure S1: Class sizes under relaxed and stricter control criteria. (A) Refractory class and (B) clustered class, recomputed with the likelihood-ratio margin over both twin statistics set to 4x, 3x (this paper), or 2x, and with all three controls or any two of the three required; filled bars are genes in the class reported in the paper, hatched bars are genes added by the relaxation, and the dotted line marks the reported class size. Clustered candidates are the 1,060 two-state genes. Every relaxation contains both reported classes in full; the 4x margin retains 127 of 186 and 235 of 329. Without any control, 1,231 genes exceed the refractory bar and 3,737 are significant in the exact test at 5 percent FDR; 744 of the 1,060 two-state genes exceed the clustered bar, of which 4 were never twin-tested and are counted in no variant. The median margin of the reported classes over their twin statistics is 5.1x (refractory) and 5.2x (clustered).

References

[1] Ido Golding, Johan Paulsson, Scott M. Zawilski, and Edward C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005. doi: 10.1016/j.cell.2005.09.031.

[2] Jonathan R. Chubb, Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. Transcriptional pulsing of a developmental gene. *Current Biology*, 16(10):1018–1025, 2006. doi: 10.1016/j.cub.2006.03.092.

[3] Arjun Raj, Charles S. Peskin, Daniel Tranchina, Diana Y. Vargas, and Sanjay Tyagi. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology*, 4:e309, 2006.

[4] Daniel Zenklusen, Daniel R Larson, and Robert H Singer. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nature Structural & Molecular Biology*, 15(12):1263–1271, 2008. doi: 10.1038/nsmb.1514.

[5] David M. Suter, Nacho Molina, David Gatfield, Kim Schneider, Ueli Schibler, and Felix Naef. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332(6028): 472–474, 2011. doi: 10.1126/science.1198817.

[6] Anton J. M. Larsson, Per Johnsson, Michael Hagemann-Jensen, Leonard Hartmanis, Omid R. Faridani, Björn Reinius, Åsa Segerstolpe, Chloe M. Rivera, Bing Ren, and Rickard Sandberg. Genomic encoding of transcriptional burst kinetics. *Nature*, 565(7738):251–254, 2019.

[7] Varun Sood and Tom Misteli. What makes genes burst. *Trends in Cell Biology*, 36(1):71–82, 2026. doi: 10.1016/j.tcb.2025.08.004.

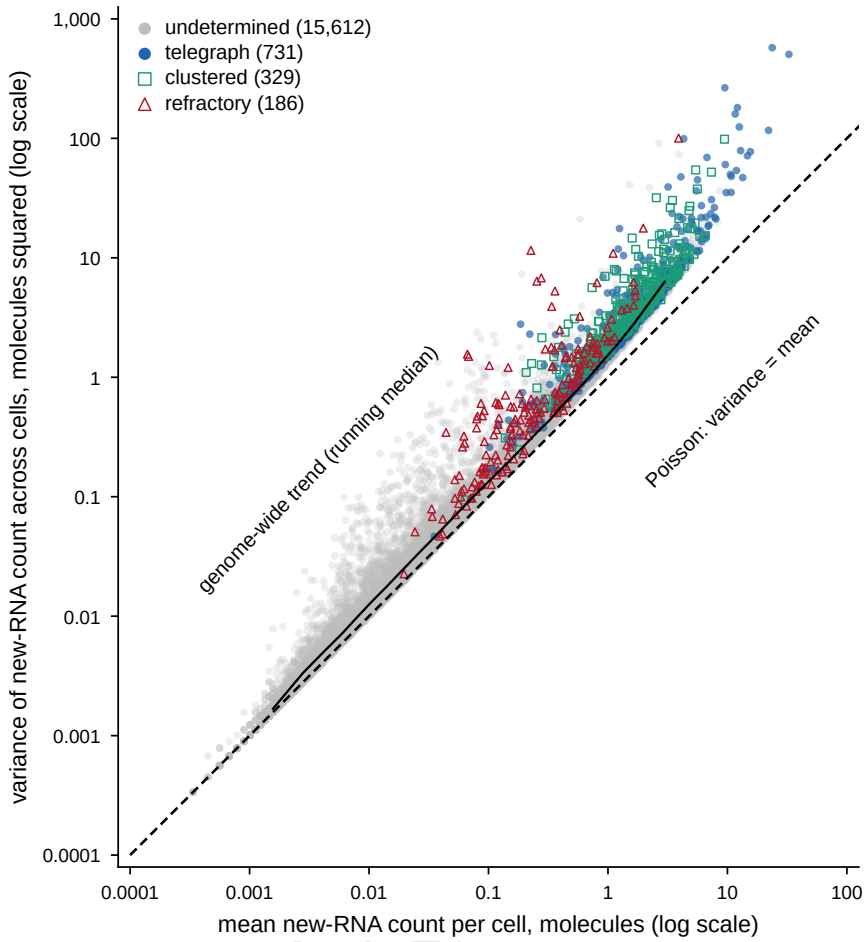

Figure S2: Cell-to-cell variance against mean of new-RNA counts for all 16,858 evaluable genes. Both axes logarithmic; points colored by class (n = 186, 329, 731 and 15,612; the 329 clustered genes are shown separately from the 731 telegraph genes); the dashed line is variance equal to mean (Poisson). Every classed gene lies above the line. Refractory and clustered genes lie above the genome-wide variance trend at matched mean (1.51-fold and 1.37-fold at the median; 91 and 92 percent of genes above the trend), whereas telegraph genes sit on it (1.06-fold; 68 percent above). Counts are the raw new-UMI counts the fits use, pooled over all 8,916 cells without depth normalization.

[8] J. Peccoud and B. Ycart. Markovian modeling of gene-product synthesis. *Theoretical Population Biology*, 48:222–234, 1995.

[9] Jong Kyoung Kim and John C. Marioni. Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biology*, 14(1):R7, 2013. doi: 10.1186/gb-2013-14-1-r7.

[10] Trung Nghia Vu, Quin F. Wills, Krishna R. Kalari, Nifang Niu, Liewei Wang, Mattias Rantalainen, and Yudi Pawitan. Beta-Poisson model for single-cell RNA-seq data analyses. *Bioinformatics*, 32 (14):2128–2135, 2016. doi: 10.1093/bioinformatics/btw202.

[11] Y. Jiang, N. R. Zhang, and M. Li. SCALE: modeling allele-specific gene expression by single-cell RNA sequencing. *Genome Biology*, 18(1):74, 2017. doi: 10.1186/s13059-017-1200-8.

[12] Hiroshi Ochiai et al. Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Science Advances*, 6:eaaz6699, 2020.

[13] Songhao Luo, Zhenquan Zhang, Zihao Wang, Xiyan Yang, Xiaoxuan Chen, Tianshou Zhou, and Jiajun Zhang. Inferring transcriptional bursting kinetics from single-cell snapshot data using a generalized telegraph model. *Royal Society Open Science*, 10(4):221057, 2023. doi: 10.1098/rsos.

221057.

[14] Wenhao Tang, Andreas Christ Sølvsten Jørgensen, Samuel Marguerat, Philipp Thomas, and Vahid Shahrezaei. Modelling capture efficiency of single-cell RNA-sequencing data improves inference of transcriptome-wide burst kinetics. *Bioinformatics*, 39(7):btad395, 2023. doi: 10.1093/bioinformatics/btad395.

[15] Florian Erhard, Marisa A. P. Baptista, Tobias Krammer, Thomas Hennig, Marius Lange, Panagiota Arampatzi, Christopher S. Jürges, Fabian J. Theis, Antoine-Emmanuel Saliba, and Lars Dölken. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature*, 571 (7765):419–423, 2019. doi: 10.1038/s41586-019-1369-y.

[16] Gert-Jan Hendriks, Lisa A. Jung, Anton J. M. Larsson, Michael Lidschreiber, Oscar Andersson Forsman, Katja Lidschreiber, Patrick Cramer, and Rickard Sandberg. NASC-seq monitors RNA synthesis in single cells. *Nature Communications*, 10(1):3138, 2019. doi: 10.1038/s41467-019-11028-9.

[17] Junyue Cao, Wei Zhou, Frank Steemers, Cole Trapnell, and Jay Shendure. Sci-fate characterizes the dynamics of gene expression in single cells. *Nature Biotechnology*, 38(8):980–988, 2020. doi: 10.1038/s41587-020-0480-9.

[18] Qi Qiu, Peng Hu, Xiaojie Qiu, Kiya W. Govek, Pablo G. Cámara, and Hao Wu. Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. *Nature Methods*, 17(10):991–1001, 2020. doi: 10.1038/s41592-020-0935-4.

[19] Daniel Ramsköld, Gert-Jan Hendriks, Anton J. M. Larsson, et al. Single-cell new RNA sequencing reveals principles of transcription at the resolution of individual bursts. *Nature Cell Biology*, 2024.

[20] Pawel Trzaskoma et al. 3D chromatin architecture, BRD4, and Mediator have distinct roles in regulating genome-wide transcriptional bursting and gene network. *Science Advances*, 10:eadl4893, 2024.

[21] Claire V. Harper, Bärbel Finkenstädt, Dan J. Woodcock, Sönke Friedrichsen, Sabrina Semprini, Louise Ashall, David G. Spiller, John J. Mullins, David A. Rand, Julian R. E. Davis, and Michael R. H. White. Dynamic analysis of stochastic transcription cycles. *PLoS Biology*, 9(4):e1000607, 2011. doi: 10.1371/journal.pbio.1000607.

[22] Benjamin Zoller, Damien Nicolas, Nacho Molina, and Felix Naef. Structure of silent transcription intervals and noise characteristics of mammalian genes. *Molecular Systems Biology*, 11:823, 2015.

[23] Joseph Rodriguez, Gang Ren, Christopher R. Day, Keji Zhao, Carson C. Chow, and Daniel R. Larson. Intrinsic dynamics of a human gene reveal the basis of expression heterogeneity. *Cell*, 176: 213–226.e18, 2019.

[24] J. Bagnall, W. Rowe, N. Alachkar, J. Roberts, H. England, C. Clark, M. Platt, D. A. Jackson, M. Muldoon, and P. Paszek. Gene-specific linear trends constrain transcriptional variability of the Toll-like receptor signaling. *Cell Systems*, 11(3):300–314.e8, 2020. doi: 10.1016/j.cels.2020.08.007.

[25] Katjana Tantale, Florian Mueller, Alja Kozulic-Pirher, Annick Lesne, Jean-Marc Victor, Marie-Cécile Robert, Serena Capozi, Racha Chouaib, Volker Bäcker, Julio Mateos-Langerak, Xavier Darzacq, Christophe Zimmer, Eugenia Basyuk, and Edouard Bertrand. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nature Communications*, 7:12248, 2016. doi: 10.1038/ncomms12248.

[26] J. Tünnermann, G. Roth, J. Cramard, and L. Giorgetti. Enhancer control of promoter activity and variability via frequency modulation of clustered transcriptional bursts. *Nature Genetics*, 58 (8):1967–1975, 2026. doi: 10.1038/s41588-026-02676-x.

[27] Andrew G. Nicoll, Juraj Szavits-Nossan, Martin R. Evans, and Ramon Grima. Transient power-law behaviour following induction distinguishes between competing models of stochastic gene expression. *Nature Communications*, 16:2833, 2025. doi: 10.1038/s41467-025-58127-4.

[28] Matthew D. Schwartz. Controlled inference of transcriptional bursting from single-cell RNA counts. Manuscript, 2026.

[29] A. R. Stinchcombe, C. S. Peskin, and D. Tranchina. Population density approach for discrete mRNA distributions in generalized switching models for stochastic gene expression. *Physical Review E*, 85(6):061919, 2012. doi: 10.1103/PhysRevE.85.061919.

[30] Joseph Rodriguez and Daniel R. Larson. Transcription in living cells: Molecular mechanisms of bursting. *Annual Review of Biochemistry*, 89:189–212, 2020. doi: 10.1146/annurev-biochem-011520-105250.

[31] Mihails Delmans and Martin Hemberg. Discrete distributional differential expression (D3E) — a tool for gene expression analysis of single-cell RNA-seq data. *BMC Bioinformatics*, 17:110, 2016. doi: 10.1186/s12859-016-0944-6.

[32] B. Reinius, J. E. Mold, D. Ramsköld, Q. Deng, P. Johnsson, J. Michaëlsson, J. Frisén, and R. Sandberg. Analysis of allelic expression patterns in clonal somatic cells by single-cell RNA-seq. *Nature Genetics*, 48(11):1430–1435, 2016. doi: 10.1038/ng.3678.

[33] V. Savova, S. Chun, M. Sohail, R. B. McCole, R. Witwicki, L. Gai, T. L. Lenz, C.-t. Wu, S. R. Sunyaev, and A. A. Gimelbrant. Genes with monoallelic expression contribute disproportionately to genetic diversity in humans. *Nature Genetics*, 48(3):231–237, 2016. doi: 10.1038/ng.3493.

[34] A. Nag, V. Savova, H. L. Fung, A. Miron, G. C. Yuan, K. Zhang, and A. A. Gimelbrant. Chromatin signature of widespread monoallelic expression. *eLife*, 2:e01256, 2013. doi: 10.7554/eLife.01256.

[35] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57:289–300, 1995.

[36] J. Kim, S. Kim, S. Jang, S. J. Park, S. Song, K. Jeung, G. Y. Jung, J.-H. Kim, H. R. Koh, and J. Sung. Cellular chemical dynamics governing signal transduction and adaptive gene expression: beyond classical kinetics. *bioRxiv*, 2026. doi: 10.64898/2026.02.13.705865.

[37] Juan M. Pedraza and Johan Paulsson. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*, 319(5861):339–343, 2008. doi: 10.1126/science.1144331.

[38] Qian Qin, Jingyu Fan, Rongbin Zheng, Changxin Wan, Shenglin Mei, Qiu Wu, Hanfei Sun, Myles Brown, Jing Zhang, Clifford A. Meyer, and X. Shirley Liu. Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biology*, 21:32, 2020.

[39] Moon Kyoo Jang, Kazuki Mochizuki, Meisheng Zhou, Ho-Sang Jeong, John N. Brady, and Keiko Ozato. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Molecular Cell*, 19(4):523–534, 2005. doi: 10.1016/j.molcel.2005.06.027.

[40] Zhiyuan Yang, Jasper H. N. Yik, Ruichuan Chen, Nanhai He, Moon Kyoo Jang, Keiko Ozato, and Qiang Zhou. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Molecular Cell*, 19(4):535–545, 2005. doi: 10.1016/j.molcel.2005.06.029.

[41] Peter B. Rahl, Charles Y. Lin, Amy C. Seila, Ryan A. Flynn, Scott McCuine, Christopher B. Burge, Phillip A. Sharp, and Richard A. Young. c-Myc regulates transcriptional pause release. *Cell*, 141(3):432–445, 2010. doi: 10.1016/j.cell.2010.03.030.

[42] Caroline R. Bartman et al. Transcriptional burst initiation and polymerase pause release are key control points of transcriptional regulation. *Molecular Cell*, 73:519–532, 2019.

[43] A. Mukherjee and M. J. Guertin. Genome-wide dynamic nascent transcript profiles reveal that most paused RNA polymerases terminate. *Nucleic Acids Research*, 54(4):gkag128, 2026.

[44] Wanqing Shao and Julia Zeitlinger. Paused RNA polymerase II inhibits new transcriptional initiation. *Nature Genetics*, 49:1045–1051, 2017.

[45] R. Vazquez Nunez, A. Bressin, Z. Shao, A. Mayer, and S. M. Vos. ELOF1 is a core component of the promoter-proximal paused RNA polymerase II complex. *bioRxiv*, 2026.

[46] Dig Bijay Mahat, H. Hans Salamanca, Fabiana M. Duarte, Charles G. Danko, and John T. Lis. Mammalian heat shock response and mechanisms underlying its genome-wide transcriptional regulation. *Molecular Cell*, 62(1):63–78, 2016. doi: 10.1016/j.molcel.2016.02.025.

[47] Anniina Vihervaara, Dig Bijay Mahat, Samu V. Himanen, Malin A. H. Blom, John T. Lis, and Lea Sistonen. Stress-induced transcriptional memory accelerates promoter-proximal pause release and decelerates termination over mitotic divisions. *Molecular Cell*, 81(8):1715–1731.e6, 2021. doi: 10.1016/j.molcel.2021.03.007.

[48] U. Lee, C. Li, C. B. Langer, N. Svetec, and L. Zhao. Comparative single-cell analysis of transcriptional bursting reveals the role of genome organization in de novo transcript origination. *Proceedings of the National Academy of Sciences*, 122(18):e2425618122, 2025. doi: 10.1073/pnas.2425618122.

[49] M. Muckenhuber, I. Seufert, K. Müller-Ott, J.-P. Mallm, L. C. Klett, C. Knotz, J. Hechler, N. Kepper, F. Erdel, and K. Rippe. Epigenetic signals that direct cell type-specific interferon beta response in mouse cells. *Life Science Alliance*, 6(4):e202201823, 2023. PMID 36732019.

[50] S. P. Burr, K. Auckland, A. Glynos, et al. MitoPerturb-Seq identifies gene-specific single-cell responses to mitochondrial DNA depletion and heteroplasmy. *Nature Structural & Molecular Biology*, 33:711–723, 2026. PMID 41922875.

[51] Barak A. Cohen, Robi D. Mitra, Jason D. Hughes, and George M. Church. A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nature Genetics*, 26(2):183–186, 2000. doi: 10.1038/79896.

[52] Huib Caron, Barbera van Schaik, Merlijn van der Mee, Frank Baas, Gregory Riggins, Peter van Sluis, Marie-Christine Hermus, Ronald van Asperen, Kathy Boon, P. A. Voûte, Siem Heisterkamp, Antoine van Kampen, and Rogier Versteeg. The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science*, 291(5507):1289–1292, 2001. doi: 10.1126/science.1056794.

[53] Martin J. Lercher, Araxi O. Urrutia, and Laurence D. Hurst. Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nature Genetics*, 31(2):180–183, 2002. doi: 10.1038/ng887.

[54] Laurence D Hurst, Csaba Pál, and Martin J Lercher. The evolutionary dynamics of eukaryotic gene order. *Nature Reviews Genetics*, 5(4):299–310, 2004. doi: 10.1038/nrg1319.

[55] Mengyi Sun and Jianzhi Zhang. Chromosome-wide co-fluctuation of stochastic gene expression in mammalian cells. *PLoS Genetics*, 15(9):e1008389, 2019. doi: 10.1371/journal.pgen.1008389.

[56] Jean Fan, Hae-Ock Lee, Soohyun Lee, Da-Eun Ryu, Semin Lee, Catherine Xue, Seok Jin Kim, Kihyun Kim, Nikolaos Barkas, Peter J. Park, Woong-Yang Park, and Peter V. Kharchenko. Linking transcriptional and genetic tumor heterogeneity through allele analysis of single-cell RNA-seq data. *Genome Research*, 28(8):1217–1227, 2018. doi: 10.1101/gr.228080.117.

[57] Richard A. Woo and Randy Y. C. Poon. Activated oncogenes promote and cooperate with chromosomal instability for neoplastic transformation. *Genes & Development*, 18(11):1317–1330, 2004. doi: 10.1101/gad.1165204.

[58] G. C. Elliott and R. T. Johnson. DNA repair in mouse embryo fibroblasts. I. Decline in ultraviolet-induced incision rate accompanies cell transformation. *Journal of Cell Science*, 60(1):267–288, 1983. doi: 10.1242/jcs.60.1.267.

[59] Simão Teixeira da Rocha, Carol A. Edwards, Mitsuteru Ito, Tsutomu Ogata, and Anne C. Ferguson-Smith. Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends in Genetics*, 24(6):306–316, 2008. doi: 10.1016/j.tig.2008.03.011.

[60] Robert D. Nicholls and Jessica L. Knepper. Genome organization, function, and imprinting in

Prader-Willi and Angelman syndromes. *Annual Review of Genomics and Human Genetics*, 2: 153–175, 2001. doi: 10.1146/annurev.genom.2.1.153.

[61] Marshall J Levesque and Arjun Raj. Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation. *Nature Methods*, 10(3):246–248, 2013. doi: 10.1038/nmeth.2372.

[62] Christopher H Bohrer and Daniel R Larson. Synthetic analysis of chromatin tracing and live-cell imaging indicates pervasive spatial coupling between genes. *eLife*, 12, 2023. doi: 10.7554/eLife.81861.

[63] Peng Dong, Shu Zhang, Valentina Gandin, Liangqi Xie, Lihua Wang, Andrew L Lemire, Wenhong Li, Hideo Otsuna, Takashi Kawase, Arthur D Lander, Howard Y Chang, and Zhe J Liu. Cohesin prevents cross-domain gene coactivation. *Nature Genetics*, 56(8):1654–1664, 2024. doi: 10.1038/s41588-024-01852-1.

[64] Heta P Patel, Stefano Coppola, Wim Pomp, Umberto Aiello, Ineke Brouwer, Domenico Libri, and Tineke L Lenstra. DNA supercoiling restricts the transcriptional bursting of neighboring eukaryotic genes. *Molecular Cell*, 83(10):1573–1587.e8, 2023. doi: 10.1016/j.molcel.2023.04.015.

[65] Jiajun Zhang, Luonan Chen, and Tianshou Zhou. Analytical distribution and tunability of noise in a model of promoter progress. *Biophysical Journal*, 102(6):1247–1257, 2012. doi: 10.1016/j.bpj.2012.02.001.

[66] M. Sinzger, M. Gehri, and H. Koeppl. Poisson channel with binary Markov input and average sojourn time constraint. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 2873–2878, 2020. doi: 10.1109/ISIT44484.2020.9174360.

[67] M. Gehri, N. Engelmann, and H. Koeppl. Mutual information of a class of Poisson-type channels using Markov renewal theory. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1931–1936, 2024. doi: 10.1109/ISIT57864.2024.10619620.

[68] A. N. Shiryaev. On optimum methods in quickest detection problems. *Theory of Probability and its Applications*, 8:22–46, 1963.

[69] S. W. Roberts. A comparison of some control chart procedures. *Technometrics*, 8:411–430, 1966.

[70] G. Lorden. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, 42(6):1897–1908, 1971. doi: 10.1214/aoms/1177693055.

[71] S. Koyama and L. Kostal. The effect of interspike interval statistics on the information gain under the rate coding hypothesis. *Mathematical Biosciences and Engineering*, 11(1):63–80, 2014. doi: 10.3934/mbe.2014.11.63.

[72] R. Ratnam and M. E. Nelson. Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. *Journal of Neuroscience*, 20(17):6672–6683, 2000. doi: 10.1523/JNEUROSCI.20-17-06672.2000.

[73] Congxin Li, Florian Cesbron, Michael Oehler, Michael Brunner, and Thomas Höfer. Frequency modulation of transcriptional bursting enables sensitive and rapid gene regulation. *Cell Systems*, 6(4):409–423.e11, 2018. doi: 10.1016/j.cels.2018.01.012.

[74] Nicholas C. Lammers, Avi I. Flamholz, and Hernan G. Garcia. Competing constraints shape the nonequilibrium limits of cellular decision-making. *Proceedings of the National Academy of Sciences*, 120(10):e2211203120, 2023. doi: 10.1073/pnas.2211203120.

[75] J. M. Kim and D. R. Larson. Timing is everything: transcription bursting in development. *Genes & Development*, 39(17–18):995–1011, 2025. doi: 10.1101/gad.352465.124.

[76] Georg Rieckh and Gašper Tkačik. Noise and information transmission in promoters with multiple internal states. *Biophysical Journal*, 106(5):1194–1204, 2014. doi: 10.1016/j.bpj.2014.01.014.

[77] E. Tunnacliffe and J. R. Chubb. What is a transcriptional burst? *Trends in Genetics*, 36(4): 288–297, 2020. doi: 10.1016/j.tig.2020.01.003.

[78] O. Shoval, H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and

U. Alon. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160, 2012.

[79] Morten Mørup and Lars Kai Hansen. Archetypal analysis for machine learning and data mining. *Neurocomputing*, 80:54–63, 2012.

[80] J. A. Redondo, T. C. Nguyen, N. Alachkar, J. Moran, A. Srivastav, M. Kochanczyk, M. Muldoon, I. Roberts, A. Krawczyk-Balska, A. Singh, and P. Paszek. Heritable single-cell gene expression states shape functional variability in innate immune responses. *bioRxiv*, 2026. doi: 10.64898/2026.06.17.732820.

[81] Michael Weber, Ines Hellmann, Michael B. Stadler, Liliana Ramos, Svante Pääbo, Michael Rebhan, and Dirk Schübeler. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*, 39:457–466, 2007. doi: 10.1038/ng1990.

[82] Tarjei S. Mikkelsen, Manching Ku, David B. Jaffe, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448:553–560, 2007. doi: 10.1038/nature06008.

[83] Vikas Rishi, Paramita Bhattacharya, Raghunath Chatterjee, Julian Rozenberg, Jianfei Zhao, Kimberly Glass, Peter Fitzgerald, and Charles Vinson. CpG methylation of half-CRE sequences creates C/EBPα binding sites that activate some tissue-specific genes. *Proceedings of the National Academy of Sciences*, 107:20311–20316, 2010. doi: 10.1073/pnas.1008688107.

[84] Jessica Charlet, Christopher E. Duymich, Fides D. Lay, Kamilla Mundbjerg, Karina Dalsgaard Sørensen, Gangning Liang, and Peter A. Jones. Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers. *Molecular Cell*, 62:422–431, 2016. doi: 10.1016/j.molcel.2016.03.033.

[85] Lacramioara Bintu et al. Dynamics of epigenetic regulation at the single-cell level. *Science*, 351: 720–724, 2016.

[86] Nathaniel A. Hathaway, Oliver Bell, Courtney Hodges, Erik L. Miller, Dana S. Neel, and Gerald R. Crabtree. Dynamics and memory of heterochromatin in living cells. *Cell*, 149(7):1447–1460, 2012.

[87] Y. Ma, M. W. Budde, J. Zhu, and M. B. Elowitz. Tuning methylation-dependent silencing dynamics by synthetic modulation of CpG density. *ACS Synthetic Biology*, 12(9):2536–2545, 2023. doi: 10.1021/acssynbio.3c00078.

[88] R. Czapiewski, M. Chiang, J. Ding, C. Naughton, G. R. Grimes, D. Marenduzzo, and N. Gilbert. Cross-chemistry single-nucleus RNA-seq identifies gene length and CpG-island promoters as determinants of transcriptional noise. *bioRxiv*, 2026. doi: 10.64898/2026.08.25.747056.

[89] S. Zhang, H. Bao, and H. Xu. Extrinsic noise or intrinsic coupling: dissecting correlated fluctuations in gene transcription. *Physical Review Letters*, 135(21):218401, 2025. doi: 10.1103/m345-slk1. bioRxiv 10.1101/2025.03.17.643613.

[90] A. Kalo, I. Kanter, A. Shraga, J. Sheinberger, H. Tzemach, N. Kinor, R. H. Singer, T. Lionnet, and Y. Shav-Tal. Cellular levels of signaling factors are sensed by β-actin alleles to modulate transcriptional pulse intensity. *Cell Reports*, 11(3):419–432, 2015. doi: 10.1016/j.celrep.2015.03.039.

[91] Qiaolin Deng, Daniel Ramsköld, Björn Reinius, and Rickard Sandberg. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*, 343(6167):193–196, 2014. doi: 10.1126/science.1245316.

[92] K. Choi, N. Raghupathy, and G. A. Churchill. A Bayesian mixture model for the analysis of allelic expression in single cells. *Nature Communications*, 10(1):5188, 2019. doi: 10.1038/s41467-019-13099-0.

[93] J. K. Kim, A. A. Kolodziejczyk, T. Ilicic, S. A. Teichmann, and J. C. Marioni. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nature Communications*, 6(1):8687, 2015. doi: 10.1038/ncomms9687.

[94] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002. doi: 10.1126/science.1070919.

[95] P. S. Swain, M. B. Elowitz, and E. D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002. doi: 10.1073/pnas.162041399.

[96] Jonathan M. Raser and Erin K. O'Shea. Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678):1811–1814, 2004. doi: 10.1126/science.1098641.

[97] Hiroshi Ochiai, Takeshi Sugawara, Tetsushi Sakuma, and Takashi Yamamoto. Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Scientific Reports*, 4:7125, 2014. doi: 10.1038/srep07125.

[98] Samuel O. Skinner, Heng Xu, Sonal Nagarkar-Jaiswal, Pablo R. Freire, Thomas P. Zwaka, and Ido Golding. Single-cell analysis of transcription kinetics across the cell cycle. *eLife*, 5:e12175, 2016. doi: 10.7554/eLife.12175.

[99] Christoph Fritzsch, Stephan Baumgärtner, Monika Kuban, Daria Steinshorn, George Reid, and Stefan Legewie. Estrogen-dependent control and cell-to-cell variability of transcriptional bursting. *Molecular Systems Biology*, 14(2):e7678, 2018. doi: 10.15252/msb.20177678.

[100] Mengyi Sun and Jianzhi Zhang. Allele-specific single-cell RNA sequencing reveals different architectures of intrinsic and extrinsic gene expression noises. *Nucleic Acids Research*, 48(2):533–547, 2020. doi: 10.1093/nar/gkz1134.

[101] Hemant Chandru Naik, Kishore Hari, Deepshikha Chandel, Susmita Mandal, Mohit Kumar Jolly, and Srimonta Gayen. Semicoordinated allelic-bursting shape dynamic random monoallelic expression in pregastrulation embryos. *iScience*, 24(9):102954, 2021. doi: 10.1016/j.isci.2021.102954.

[102] Parichitran Ayyamperumal, Hemant Chandru Naik, Amlan Jyoti Naskar, Lakshmi Sowjanya Bammidi, and Srimonta Gayen. Epigenomic states contribute to coordinated allelic transcriptional bursting in iPSC reprogramming. *Life Science Alliance*, 7(4):e202302337, 2024. doi: 10.26508/lsa.202302337.

[103] Jeff E. Mold, Martin H. Weissman, Michael Ratz, Michael Hagemann-Jensen, Joanna Hård, Carl-Johan Eriksson, Hosein Toosi, Joseph Berghenstråhle, Christoph Ziegenhain, Leonie von Berlin, Marcel Martin, Kim Blom, Jens Lagergren, Joakim Lundeberg, Rickard Sandberg, Jakob Michaëlsson, and Jonas Frisén. Clonally heritable gene expression imparts a layer of diversity within cell types. *Cell Systems*, 15(2):149–165.e10, 2024. doi: 10.1016/j.cels.2024.01.004.

[104] L. Marion-Poll, B. Forêt, D. Zielinski, F. Massip, M. Attia, A. C. Carter, L. Syx, H. Y. Chang, A.-V. Gendrel, and E. Heard. Locus specific epigenetic modalities of random allelic expression imbalance. *Nature Communications*, 12(1):5330, 2021. doi: 10.1038/s41467-021-25630-3.

[105] O. Stewart, C. Gruber, Haley E. Randolph, et al. Monoallelic expression can govern penetrance of inborn errors of immunity. *Nature*, 2025. doi: 10.1038/s41586-024-08346-4.

[106] Michael B. Heskett, Athanasios E. Vouzas, Brian Johnstone, et al. Autosomal allelic inactivation at loci with variable replication timing and dosage sensitivity. *eLife*, 2026. doi: 10.7554/eLife.109938.

[107] D. S. Foster, M. Januszyk, K. E. Yost, M. S. Chinta, G. S. Gulati, A. T. Nguyen, A. R. Burcham, A. Salhotra, R. C. Ransom, D. Henn, K. Chen, S. Mascharak, K. Tolentino, A. L. Titan, R. E. Jones, O. da Silva, W. T. Leavitt, C. D. Marshall, H. E. des Jardins-Park, M. S. Hu, D. C. Wan, G. Wernig, D. Wagh, J. Coller, J. A. Norton, G. C. Gurtner, A. M. Newman, H. Y. Chang, and M. T. Longaker. Integrated spatial multiomics reveals fibroblast fate during tissue repair. *Proceedings of the National Academy of Sciences of the United States of America*, 118:e2110025118, 2021. doi: 10.1073/pnas.2110025118. PMID 34620713; GSE178758.

[108] Jingyi Ren, Haowen Zhou, Hu Zeng, Connie Kangni Wang, et al. Spatiotemporally resolved transcriptomics reveals the subcellular RNA kinetic landscape. *Nature Methods*, 20:695–705, 2023.

[109] O. A. Stefansson, B. D. Sigurpalsdottir, S. Rognvaldsson, G. H. Halldorsson, K. Juliusson, G. Sveinbjornsson, B. Gunnarsson, D. Beyter, H. Jonsson, S. A. Gudjonsson, T. A. Olafsdottir, S. Saevarsdottir, M. K. Magnusson, S. H. Lund, V. Tragante, A. Oddsson, M. T. Hardarson, H. P. Eggertsson, R. L. Gudmundsson, S. Sverrisson, M. L. Frigge, F. Zink, H. Holm, H. Stefansson, T. Rafnar, I. Jonsdottir, P. Sulem, A. Helgason, D. F. Gudbjartsson, B. V. Halldorsson, U. Thorsteinsdottir, and K. Stefansson. The correlation between CpG methylation and gene expression is driven by sequence variants. *Nature Genetics*, 56(8):1624–1631, 2024. doi: 10.1038/s41588-024-01851-2.

[110] V. Mutzel, T. Schwämmle, S. Oeverdieck, L. Librenjak, B. Boesen, M. Bothe, R. A. F. Gjaltema, I. Dunkel, G. Noviello, and E. G. Schulz. Antisense transcription can induce expression memory via stable promoter repression. *Genome Biology*, 26(1):430, 2025. doi: 10.1186/s13059-025-03875-1.

[111] Jr. Johnston, R. J. and C. Desplan. Interchromosomal communication coordinates intrinsically stochastic expression between alleles. *Science*, 343(6171):661–665, 2014. doi: 10.1126/science.1243039.

[112] M. A. Kerlin, I. Aboulfath-Ladid, J. Roensch, C. Jaubert, A. Battistella, K. J. E. Borgman, and A. Coulon. Transcriptional coregulation in cis around a contact insulation site revealed by single-molecule microscopy. *eLife (reviewed preprint)*, 2025. doi: 10.7554/elife.106678. reviewed preprint, no version of record as of 2026-09-08; bioRxiv 10.1101/2025.03.04.641491.

[113] Constantinos Chronis, Petko Fiziev, Bernadett Papp, Stefan Butz, Giancarlo Bonora, Shan Sabri, Jason Ernst, and Kathrin Plath. Cooperative binding of transcription factors orchestrates reprogramming. *Cell*, 168(3):442–459.e20, 2017. doi: 10.1016/j.cell.2016.12.016.

[114] Zachary R. Fox and Brian Munsky. The finite state projection based Fisher information matrix approach to estimate information and optimize single-cell experiments. *PLoS Computational Biology*, 15(1):e1006365, 2019. doi: 10.1371/journal.pcbi.1006365.