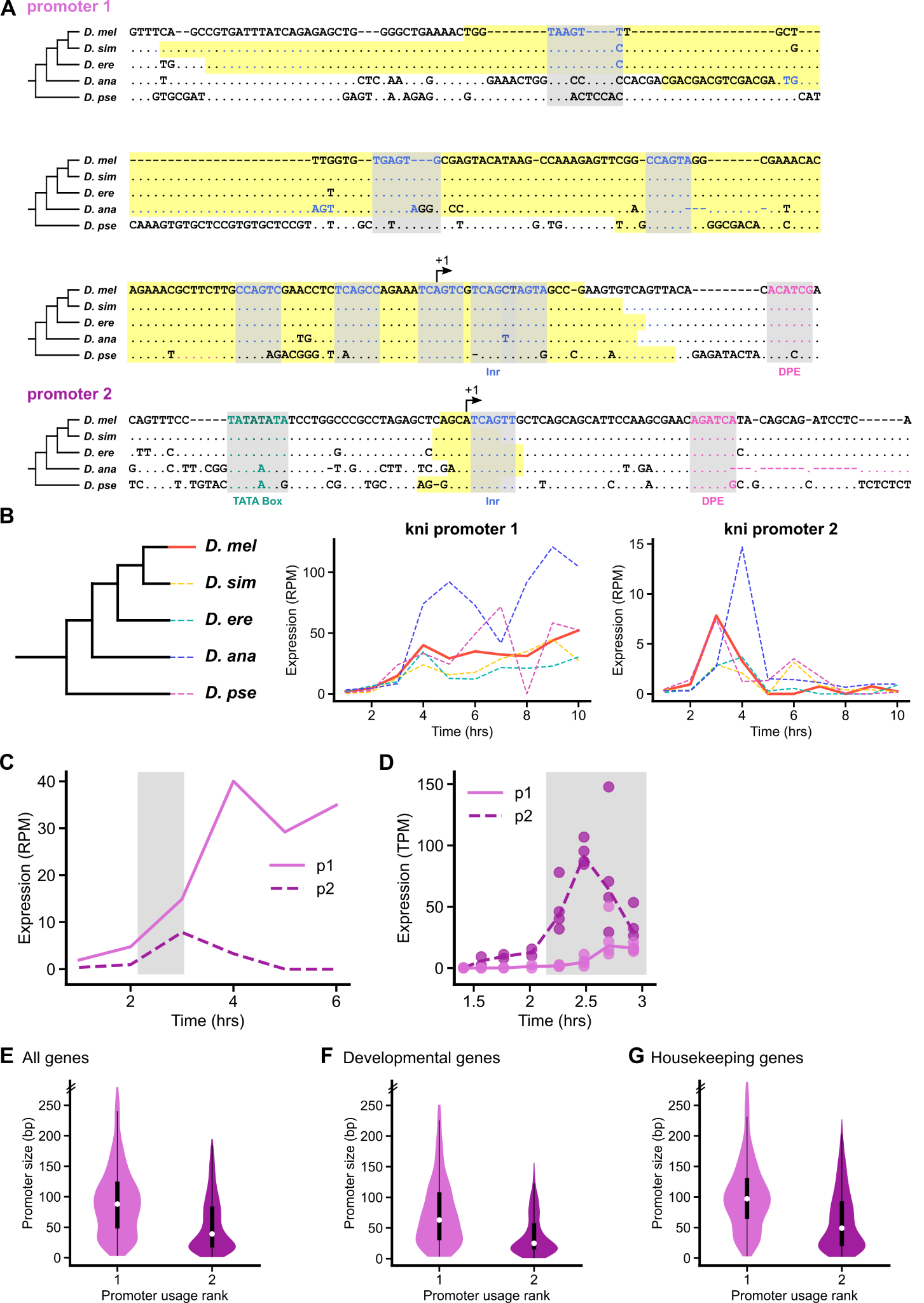
****

**Figure S1. The *knirps* promoters show sequence and functional conservation, and this two-promoter structure is prevalent among genes expressed during development.**

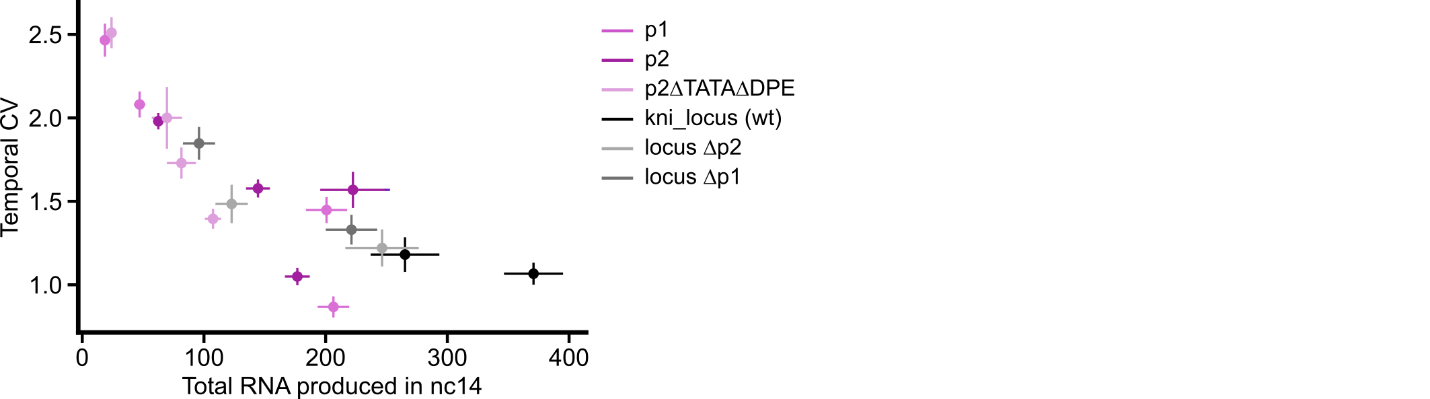
**(A)** Both *kni* promoters are aligned with the orthologous sequences in four other *Drosophila* species, with dashes (-) representing unaligned sequence and dots (.) indicating matching base pairs. There is remarkable sequence conservation, with the core promoter motifs preserved across all five species. The highlighted regions represent transcription start clusters (TSCs), identified by Batut, *et al*. (Batut and Gingeras 2017) as regions of statistically significant clustering of cDNA 5’ ends. **(B)** *Kni* promoter activity over the first 10 hours of development is reasonably consistent across five species of *Drosophila*, with promoter 1 generally being used more than promoter 2. Specifically, note that both promoters are used in nuclear cycle 14 (2-3 hours) in all five species. Expression data for the two *kni* promoters is shown, with RAMPAGE data (Batut and Gingeras 2017) in **(C)** and RNA-seq data (Lott *et al*. 2011) in **(D)**; the time period corresponding to the blastoderm stage (nuclear cycle 14) is highlighted in gray. Based on these two sets of data, the two *kni* promoters are both used during nuclear cycle 14. The discrepancy between these two sets of data may be explained by the absence of multiple replicates of RAMPAGE data, which is likely within the range of error during nuclear cycle 14. **(E – G)** For developmentally expressed genes with multiple promoters that are represented in both the Eukaryotic Promoter Database and the Batut *et al.* RAMPAGE data (Batut and Gingeras 2017; Dreos *et al.* 2017), violin plots of the two most used promoters, with the primary promoter (most used) in light purple and the secondary promoter (second most used) in purple. The black boxes span the lower to upper quartiles, with the white dot within the box indicating the median. Whiskers extend to 1.5\*IQR (interquartile range) ± the upper and lower quartile, respectively. The double hash marks on the axes indicate that 95% of the data is being shown. **(E)** When the two most used promoters of genes expressed in embryogenesis (*n* = 1177) are plotted, the size of primary promoters is significantly larger than that of the secondary promoter. **(F)** When limited to promoters of developmentally controlled genes – genes whose expression pattern varies considerably as a function of developmental time – (*n* = 387) this trend of larger primary promoters is maintained, though on average, these promoters are sharper that those of the whole gene set in panel E. **(G)** When limited to promoters of housekeeping genes (*n* = 790), this trend of larger primary promoters is also maintained, though on average, these promoters are still broader than those of developmentally controlled genes.

A picture containing timeline

Description automatically generated

**Figure S2. TFs show preferences for certain core promoter motifs.**

To identify patterns of TF-core promoter motif co-occurrence, we identified a set of enhancer-gene pairs from multiple databases that were found through enhancer-trapping experiments (Bonn *et al*., 2012; Halfon *et al.* 2008), a high-throughput tiling screen (Kvon *et al*. 2014), and 4C-seq (Ghavi-Helm *et al*. 2014) and limited the analysis to genes expressed during early development. The promoters were from the Eukaryotic Promoter Database (Dreos *et al*. 2017). These enhancer-gene pairs can be found in Supplementary File 2. We then calculated the fold enrichment over background of core promoter elements associated with TF-target genes. The left heatmap shows the log fold-enrichment over background of core promoter motifs (columns) and the TFs (rows) controlling their target genes. The right heatmap shows the log fold-enrichment over background of the motif combinations (columns) and the TF (rows) controlling their target genes. For example, this means that column 1 (Inr) in the left heatmap shows enrichment of any promoters that contain Inr regardless of any other promoter motifs they might contain, whereas column 1 (Inr only) in the right heatmap shows enrichment of promoters with only Inr and no other predicted core promoter motifs.

****

**Figure S3. Noise is inversely correlated with total RNA produced.**

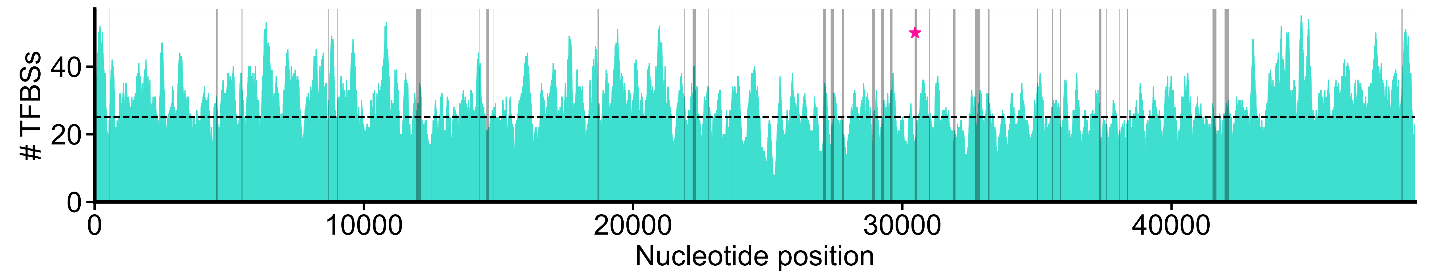
To examine the relationship between the temporal coefficient of variation (CV) and activity of each construct, we plotted the mean temporal CV against the total RNA produced in nc14 at the anterior-posterior bin of maximum expression (22% and 63%) for the anterior band and the posterior stripe, respectively, with the error bars representing 95% confidence intervals. There is a clear trend of CV decreasing with increased total RNA produced though there are examples where constructs with the same promoter can produce higher noise than others with similar output levels, suggesting that promoters do not solely dictate noise levels.

**A screenshot of a computer

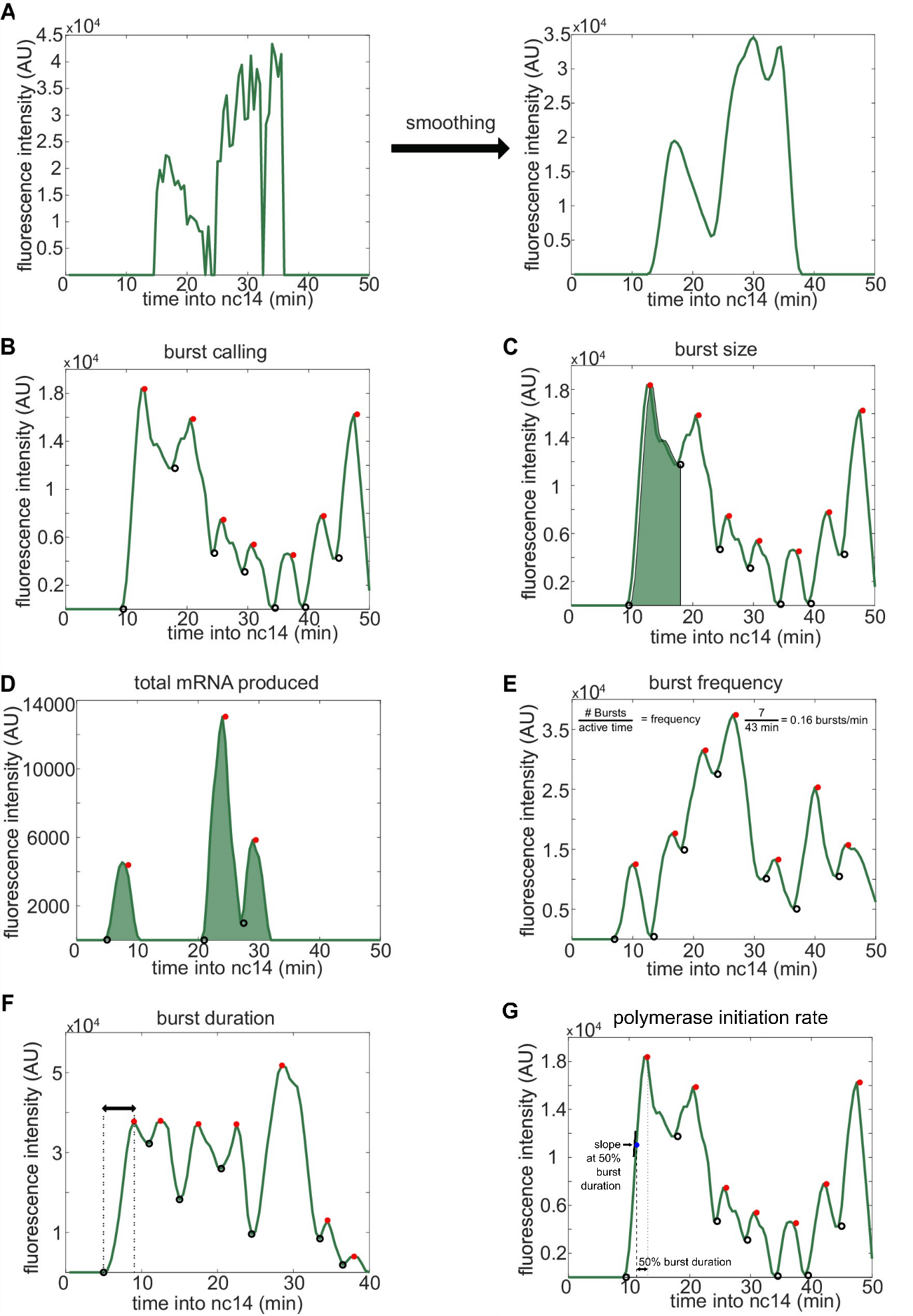
Description automatically generated with medium confidence**

**Figure S4. Two-state model of transcription in the context of tracking transcription dynamics.**

**(A)** In the two-state model of transcription, in which the promoter is either (1) in the inactive state (“OFF”), in which RNA polymerase cannot bind and initiate transcription or (2) in the active state (“ON”), during which it can. The promoter transitions between these two states with rates *kact* and *kinact*, with promoter activation involving both the interaction of the enhancer and promoter and the assembly of all the necessary transcription machinery for transcription initiation to occur. This may occur through enhancer looping or through the formation of a transcriptional hub. In its active state, the promoter produces mRNA at rate *r*, and the mRNA decays by diffusing away from the gene locus at rate *µ*. **(B)** MS2-tagging RNA allows us to track nascent transcription, and the resulting fluorescence trace (in light blue) is proportional to the number of nascent RNA produced over time. The graph is split into sections, representing different molecular states and how they correspond to our transcription measurements. These states are represented by different colors—red when the promoter is OFF, green when it is ON, and yellow when transcription continues but the promoter is no longer ON, as no new polymerases are being loaded. The dynamics of these fluctuations or bursts can be characterized by quantifying various properties, including burst frequency (how often a burst a occurs), burst size (number of RNA produced per burst), and burst duration (the period of active transcription during which mRNA is produced at rate *r*).

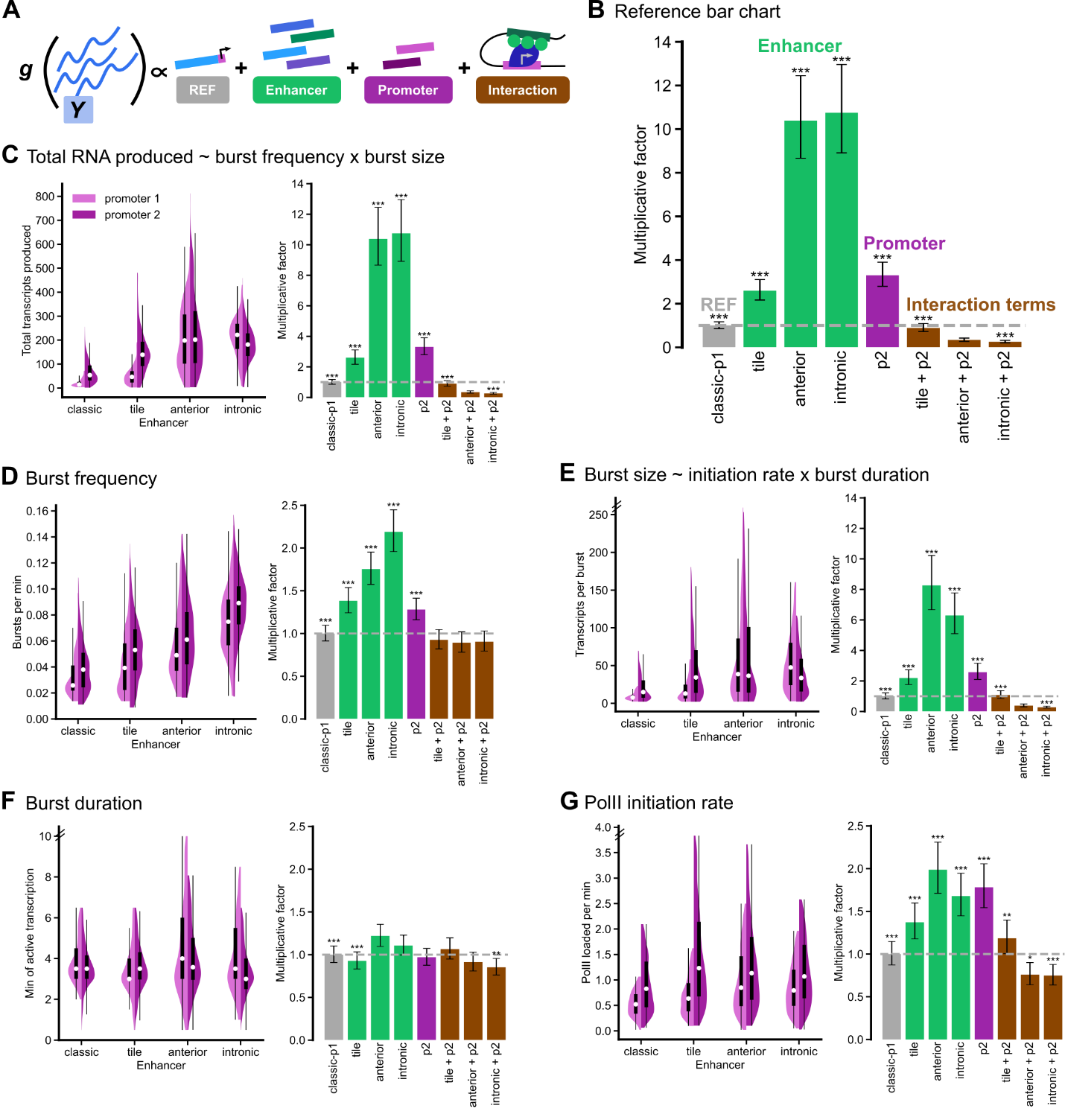
**Figure S5. The lambda genome analyzed for regions with no core promoter motifs and minimal TF binding sites.**

The number of predicted TF binding sites in a 156 bp sliding window (i.e. the size of promoter 1) are plotted in cyan against their nucleotide position in the lambda phage genome. The shaded gray regions represent windows in which there are no predicted core promoter motifs. The dashed black line indicates the average number of TF binding sites in the regions lacking promoter motifs. The pink star represents the location of the fragment of lambda DNA that was chosen to replace p1 in the locus ∆p1 construct.



**Figure S6. Visual inspection of burst calling algorithm.**

This figure is adapted from Waymack, *et al*. (Waymack *et al*. 2020) with one additional panel **(G)** added. To quantify the burst properties of interest (burst size, burst frequency, burst duration, and polymerase initiation rate), we began by smoothing individual fluorescence traces using the LOWESS method with a span of 10%. Periods of promoter activity or inactivity were then determined based on the slope of the fluorescence trace. **(A)** Example of smoothing transcriptional traces. **(B)** Fluorescence trace of a single punctum during nc14. Open black circles indicate time points where the promoter has turned “on”, filled red circles indicate time points where the promoter is identified as turning “off”. **(C)** Transcriptional trace with the green shaded region under the curve used to calculate the size of the first burst. This area of this region is calculated using the trapz function in MATLAB and extends from the time point the promoter is called “on” until the next time it is called “on”. Panels **(D – F)** show additional representative fluorescence traces of single transcriptional puncta during nc14. **(D)** A trace with the entire region under the curved shaded green represents the area used to calculate the total amount of mRNA produced. This area is calculated using the trapz function in MATLAB extends from the time the promoter is first called “on” until 50 min into nc14 or the movie ends, whichever comes first. **(E)** Burst frequency is calculated by dividing the number of bursts that occur during nc14 by the length of time from the first time the promoter is called “on” until 50 min into nc14 or the movie ends, whichever comes first. **(F)** Burst duration is calculated by taking the amount of time between when the promoter is called “on” and it is next called “off”. **(G)** Polymerase initiation rate is calculated by taking the slope of the smoothed fluorescence race at the midpoint between when the promoter is called “on” and it is next called “off”.



**Figure S7. Using the log link function gives the same results.**

Here, we show the results from the generalized linear models (GLMs) when using the log link function instead of the identity link function, which was used in Figure 4.

**(A)** To parse the effects of the enhancer, the promoter, and their interactions on all burst properties, we built GLMs. *Y* represents the burst property under study, *g* is the link function, and the enhancers, promoters, and their interaction terms are the explanatory variables. The coefficients of each of these explanatory variables are representative of their contribution to the total value of the burst property.All burst property data was taken from the anterior-posterior bin of maximum expression (22% and 63%) for the anterior band and the posterior stripe, respectively. Note that modeling attempts to fit the means, but medians are shown in the violin plots. **(B)** The classic-p1 construct was chosen as a reference and is represented in gray, and the effects of enhancer, promoter, and their interactions are represented in green, purple, and brown, respectively. We exponentiate the coefficients and the 95% confidence intervals for each independent variable to invert the log link function and call these quantities the “multiplicative factors.” Performing this conversion yields a multiplicative relationship between our response variable (the burst property) and our explanatory variables. The reference construct (classic-p1) has been set to 1 such that multiplying the relevant multiplicative factors gives you the value that, if multiplied by the reference construct value, will gives you the average value of the burst property for a particular construct. These factors are plotted as a bar graph; \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001. In panels **(C – G)**, (left) split violin plots (and their associated box plots) of burst properties for all eight constructs will be plotted with promoter 1 in light purple and promoter 2 in purple. The black boxes span the lower to upper quartiles, with the white dot within the box indicating the median. Whiskers extend to 1.5\*IQR (interquartile range) ± the upper and lower quartile, respectively. (right) Bar graphs representing the relative contributions of enhancer, promoter, and their interactions to each burst property are plotted as described in **(B)**. The double hash marks on the axes indicate that 90% of the data is being shown. **(C)** Expression levels are significantly affected by the enhancer identity (green bars), promoter identity (purple bars) and the interaction terms (brown bars), with interaction terms representing the role of molecular compatibility. **(D)** Burst frequency is dominated by the promoter terms, with promoter 2 consistently producing higher burst frequencies regardless of enhancer. **(E)** Burst size is in part determined by initiation rate and burst duration. Burst size is significantly affected by the enhancer, promoter, and interaction terms. **(F)** Burst duration is reasonably consistent regardless of enhancer or promoter, but some enhancer and interaction terms have small, but significant effects. **(G)** Pol II initiation rate is significantly affected by enhancer, promoter, and interaction terms, with promoter 2 consistently driving higher initiation rates than promoter 1. However, the differences in promoter 1 and promoter 2 initiation rates is enhancer-dependent, as indicated by the significant interaction terms.