# Differential regulation of alternative promoters emerges from unified kinetics of enhancer-promoter interaction

Jingyao Wang[1,2], Shihe Zhang[2,3,*], Hongfang Lu[1,2], and Heng Xu[2,3,*]

[1]School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[2]Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China

[3]School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

*Correspondence: hsx_sjtu@sjtu.edu.cn (S.Z.), Heng_Xu@sjtu.edu.cn (H.X.)

**Abstract**

Many eukaryotic genes contain alternative promoters with distinct expression patterns. How these promoters are differentially regulated remains elusive. Here, we apply single-molecule imaging to quantify the transcriptional regulation of two alternative promoters (P1 and P2) of the Bicoid (Bcd) target gene *hunchback* in syncytial blastoderm *Drosophila* embryos. Contrary to the previous notion that Bcd only activates P2, we find that Bcd activates both promoters via the same two enhancers. P1 activation is less frequent and requires binding of more Bcd molecules than P2 activation. Using a theoretical model to relate promoter activity to enhancer states, we show that the two promoters follow common transcription kinetics driven by sequential Bcd binding at the two enhancers. Bcd binding at either enhancer primarily activates P2, while P1 activation relies more on Bcd binding at both enhancers. These results provide a quantitative framework for understanding the kinetic mechanisms of complex eukaryotic gene regulation.

## Introduction

Promoters initiate gene transcription by interacting with specific *cis*-regulatory sequences (enhancers)[1-3]. In eukaryotic genomes, many genes contain alternative promoters[4,5], which can produce functionally distinct transcript isoforms under the regulation of multiple enhancers[6,7]. During development, the differential expression of these isoforms is critical to cell fate decisions[8,9]. Misregulation of these isoforms can lead to diseases, including cancer[6,10-12]. Thus, eukaryotic gene regulation needs to be understood at the level of individual alternative promoters and enhancers.

Enhancer activation of a single promoter involves a series of molecular events, including transcription factor (TF) binding, local chromatin opening, and physical proximity between the two elements[1,2,13]. Typically, TF binding determines the spatial expression pattern[1,14], whereas other events set the expression amplitude[15-18]. When multiple promoters and enhancers are present, interactions between these elements often result in complex and varied promoter behaviors[19]. For example, multiple enhancers may contact each other and synergistically drive a promoter[20,21]. Alternatively, different enhancers may be mutually exclusive and compete for promoter activation[22,23]. In these cases, enhancers combine their regulatory effects differently, ranging from superadditive to subadditive or even repressive summation[12]. Similarly, a single enhancer may activate multiple promoters simultaneously[24,25] or one at a time[26]. So far, although specific mechanisms were proposed for some of these phenomena, it is unclear whether

universal mechanisms exist for the complex interactions between multiple enhancers and promoters.

An ideal model to investigate alternative promoter regulation is the *Drosophila* gap gene *hunchback* (*hb*), which contains two alternative promoters (P1 and P2, Fig. 1a)[27-30]. In early embryogenesis (syncytial blastoderm stage, nuclear division cycles (nc) 10–13), *hb* is expressed in a bursty manner[31,32] throughout the anterior half of the embryo in response to the concentration gradient of the maternal TF Bicoid (Bcd)[29,33,34]. This expression is believed to be purely from the P2 promoter mediated by two enhancers: a proximal enhancer located next to P2[29,34] and a distal "shadow" enhancer located 4.5 kb upstream[35]. Cooperative Bcd binding at either enhancer can activate P2[35]. Competitive action of these partially redundant enhancers helps suppress expression noise and ensure a robust expression pattern[23,35].

Unlike P2, the P1 promoter is believed to be inactive during early development[36,37], due to its lack of Zelda binding sites and TATA box necessary for local chromatin opening[30,38]. Instead, the promoter is activated in late nc14 by a Bcd-independent stripe enhancer[28,36]. However, previous measurements of endogenous P1 activity relied on traditional *in situ* hybridization methods, which have limited sensitivity to detect weak mRNA signals. In fact, *hb*-reporter experiments showed that P1 could respond to Bcd when placed adjacent to the proximal or distal enhancers[30]. Without precise quantification of endogenous P1 activity, the understanding of early

67 *hb* regulation is incomplete. It is unclear how P1 interacts with different *hb* enhancers and
68 whether the mechanisms of P1 and P2 regulation are intrinsically related.

69 Here, we use single-molecule fluorescence *in situ* hybridization (smFISH)[31,32,39,40] to quantify
70 the expression of individual P1 and P2 promoters for each endogenous *hb* gene locus in nc11–
71 13 embryos. Contrary to the previous notion, we find that P1 contributes a modest yet non-
72 negligible fraction of early *hb* transcription and affects the expression patterns of other gap genes.
73 Using different TF dosages and enhancer deletion, we show that Bcd activates both promoters
74 via the proximal and distal enhancers. Compared with P2, P1 activation requires cooperative
75 binding of more Bcd molecules and a synergistic (as opposed to competitive) action of the two
76 enhancers. Analyzing the statistics of nascent mRNA signals from individual promoter loci
77 reveals that both promoters follow a unified scheme of three-state transcription kinetics.
78 Cooperative Bcd binding at either enhancer can drive a promoter to a weak active state, while
79 additional Bcd binding at the second enhancer can turn the promoter to full-power transcription.
80 The two promoters differ in their responses to different Bcd binding configurations. P2
81 transcription is primarily driven by Bcd binding at a single enhancer, while P1 transcription relies
82 more on Bcd binding at both enhancers. In concert, these results provide a simple and
83 quantitative mechanism for the differential regulation of alternative promoters. Our quantitative
84 approach may be generalized as a framework for deciphering complex eukaryotic gene
85 regulation involving multiple promoters and enhancers.

## Results

### *hb* P1 and P2 promoters are both active in early embryogenesis.

Each *hb* promoter produces a unique transcript isoform (P1: *hb*-RB, P2: *hb*-RA; Supplementary Fig. 1a-c; see Supplementary Note)[29,41]. To quantify the endogenous P1 and P2 activities in early embryos, we applied smFISH with four sets of oligonucleotide probes designed for different regions of *hb* mRNAs. Specifically, two P1-specific probe sets targeted the 5' untranslated region (UTR) and intron region of *hb*-RB. A P2-specific probe set targeted the 3'UTR of *hb*-RA. Finally, a probe set targeted the coding sequence (CDS) region shared by *hb*-RA and *hb*-RB (Fig. 1a, Supplementary Table 1).

Confocal imaging and automated image analysis identified actively transcribing *hb* loci as bright FISH spots in wild-type (WT) syncytial blastoderm embryos (Fig. 1b). We quantified the instantaneous transcriptional activity of every locus in units of individual cytoplasmic mRNAs[32] (Fig. 1b; Supplementary Fig. 2; see Methods). For the CDS signal, we observed bright FISH spots in the anterior part of nc11–13 embryos (Fig. 1c). Most anterior nuclei contained two bright spots, while some nuclei exhibited three or four spots due to the replication of the *hb* gene (Supplementary Fig. 3a). These results are consistent with literature[31], indicating that at least

one of the two *hb* promoters is active in the anterior side of the embryo.

Similar to the CDS signal, promoter-specific FISH signals were also concentrated in the anterior half of nc11–13 embryos (Fig. 1c). Specifically, ~70% of the anterior nuclei (within the range of 0.20–0.40 embryo length (EL)) showed bright P2-3'UTR spots (Fig. 1d), consistent with previous reports of active P2 expression in early development[27,30]. Surprisingly, ~32%–57% of the anterior nuclei in nc11–13 embryos also contained bright P1-5'UTR and intron spots (Fig. 1d), indicating early P1 transcription, a phenomenon barely observed in previous bulk studies (Supplementary Fig. 1d, e). The percentage of P1-active anterior nuclei increased with the nuclear cycle (Fig. 1d), suggesting that P1 becomes increasingly active during development. In total, P1- and P2-specific probe signals were exhibited in ~73%–84% of the anterior nuclei, lower than that of the CDS signal. This percentage difference may be because the P2-specific probes target the very end of *hb*-RA, which is missing in many incomplete nascent transcripts.

To analyze the instantaneous transcription of P1 and P2, we plotted, for each probe set, the nascent mRNA signal against the nuclear position for all nuclei in the embryo (Fig. 1e; Supplementary Fig. 3b). During nc11–13, the average expression profile for each probe set exhibited a reverse-sigmoidal shape within the range 0.20–0.70 EL. By fitting the profile to a logistic function, we estimated the maximal signal level $r_{max}$ and the boundary position $x_0$ of the anterior expression domain (Supplementary Fig. 3b; see Methods). As expected, $r_{max}$ of P1-specific signals was much smaller than that of the CDS signal (Fig. 1f), indicating that P1

123 contributes less than P2 to early *hb* expression. $r_{max}$ of the P2-specific signal also remained low
124 (Fig. 1f) due to the probes' target location at the 3' end of *hb*-RA. $r_{max}$ of the P1- and P2-specific
125 signals should not be directly compared, since their probes target different ends of nascent
126 transcripts (5' vs. 3'; see next section for a comparison between P1- and P2-specific signal
127 levels). The boundary position $x_0$ of the P1-specific signals was significantly lower than that of
128 the P2-specific and CDS signals (~0.42 EL v.s. ~0.46EL, $p<0.05$, Student's t-test, Fig. 1e, g),
129 indicating that P1 is activated at a more anterior position than P2.

130 In addition to the mean expression level, the activities of individual promoter loci exhibited
131 substantial variability (Fig. 1b). The Fano factors of anterior P1- and P2-specific nascent signals
132 were much larger than one (Fig. 1h; see Methods), indicating bursty transcription from both
133 promoters[42,43]. Such burstiness primarily resulted from the intrinsic stochasticity of *hb*
134 transcription, as the two promoters behaved independently (Supplementary Fig. 3c, d; see
135 Methods). Quantifying the intrinsic noise[44] revealed that P1 was much noisier than P2
136 (Supplementary Fig. 3e; see Methods), consistent with its low expression level.

137

138 **P1 contributes a non-negligible fraction of early *hb* transcription and function.**

139 To quantify the contributions of P1 and P2 to early *hb* transcription, we note that the *hb* CDS
140 signal reflects either P1 or P2 transcription. Using CRISPR-based mutant fly lines with P1 and

141 P2 deletions (ΔP1C and ΔP2C)[30], we showed that the average nascent CDS signal from either
142 promoter was proportional to the signal of promoter-specific probes (Supplementary Fig. 4; see
143 Supplementary Note). Thus, the nascent CDS signal ($r_{CDS}$) in the WT embryo is a linear
144 combination of the nascent P1-5'UTR and P2-3'UTR signals ($r_{P1}$ and $r_{P2}$):

145 $$r_{CDS} = a_1 r_{P1} + a_2 r_{P2} \quad (1)$$

146 where $a_1$ and $a_2$ are ratio parameters that, in theory, depend on the mRNA elongation and
147 termination dynamics (see Supplementary Note)[45]. The two terms on the right-hand side of the
148 equation distinguish the contribution of each promoter to the CDS signal.

149 To examine Equation (1), we compared nuclear P1-5'UTR, P2-3'UTR, and CDS signals in
150 the same embryo (Fig. 2a; see Methods). The three signals satisfied a linear relationship (Fig.
151 2b) with $a_1$ = 0.53 ± 0.03 and $a_2$ = 3.72 ± 0.90 (Fig. 2c). In a simple transcription model with
152 deterministic elongation and termination processes[45,46], the above $a_1$ and $a_2$ values indicate a
153 post-elongation residence time ($T_R$) in terminating a nascent mRNA (Fig. 2d; Supplementary Fig.
154 5b; see Supplementary Note). The estimated $T_R$ for P1 and P2 was 142 s and 22 s, respectively,
155 consistent with previous estimations of transcription termination[45,47,48]. These results validate
156 Equation (1) and enable the decomposition of the nascent CDS signal into different promoter
157 activities (Supplementary Fig. 5c; see Methods). In the anterior expression domain, the
158 contribution of P1 increased with the nuclear cycle to ~13% (Fig. 2e), which was modest yet
159 non-negligible.

In addition to the mean expression level, we investigated the contributions of P1 and P2 to the intrinsic noise of nascent *hb* transcription. Considering that P1 and P2 activities were independent, their contributions to CDS noise ($\eta^2_{\text{CDS}}$) could be summed as:

$$\eta^2_{\text{CDS}} = \eta^2_{\text{P1-CDS}} f^2_{\text{P1-CDS}} + \eta^2_{\text{P2-CDS}} f^2_{\text{P2-CDS}} \quad (2)$$

where $\eta^2_{\text{P1-CDS}}$ and $\eta^2_{\text{P2-CDS}}$ denote CDS noises originating from P1 and P2 activities, respectively, while $f_{\text{P1-CDS}}$ and $f_{\text{P2-CDS}}$ indicate the relative contributions of each promoter to the mean CDS signal (see Supplementary Note). Equation (2) revealed that P1 contributed to up to ~20% of the anterior CDS noise during nc11–13 (Fig. 2f), which exceeded its contribution to the mean CDS signal.

To evaluate the function of early P1 transcription in embryo development, we measured the expression patterns of *hb*-target genes, *Krüppel* (*Kr*) and *knirps* (*kni*), in P1 deletion embryos (ΔP1C)[30]. Based on the literature, the anterior boundaries of *Kr* and *kni* expression domains in nc14 are both set by the repression of Hb protein[49,50]. Consistent with this picture, the anterior *Kr* and *kni* boundaries shifted anteriorly for ~0.09 EL (from 0.45 ± 0.02 EL in WT to 0.36 ± 0.01 EL in ΔP1C, $p<0.01$) and ~0.06 EL (from 0.58 ± 0.01 EL in WT to 0.52 ± 0.02 EL in ΔP1C, $p<0.05$), respectively, in response to P1 deletion (Fig. 2g, h). Thus, early P1 transcription has a non-negligible effect on *Drosophila* AP patterning. However, no obvious cuticle disruption was reported for ΔP1C larvae[30], suggesting compensation from other genes in the downstream patterning process.

179

180 **Activating P1 requires cooperative binding of more Bcd molecules than activating P2.**

181 In early embryogenesis, the exponential gradient of Bcd is the primary driver of anterior *hb*
182 transcription[29,34]. To examine whether Bcd directly regulates both promoters, we used a
183 transgenic fly line with 1× functional *bcd* gene[51]. With reduced Bcd dosage, the anterior
184 expression domains of P1 and P2 both retreated towards the anterior pole (Fig. 3a). Specifically,
185 P1 and P2 expression boundaries in 1× *bcd* embryos significantly shifted to the anterior side by
186 ~0.07 EL ($p<0.01$, Student's t-test, Fig. 3b). Thus, Bcd activates both promoters in nc11–13
187 embryos.

188 To quantify the regulation of each promoter by Bcd, we combined promoter-specific smFISH
189 of *hb* mRNA with immunofluorescence of Bcd (Fig. 1b). We estimated the absolute Bcd
190 concentration in each nucleus using a previously developed image analysis method[32]
191 (Supplementary Fig. 6a, b; see Methods and Supplementary Note). For each promoter, we
192 plotted the average transcriptional response per nucleus (0.25–0.75 EL) to Bcd concentration in
193 the embryo (Fig. 3c), known as the gene regulation function (GRF)[52]. Previous studies reported
194 that *hb* GRF fitted well to a Hill function with a Hill coefficient $h \approx 5$–$6$[32,53]. A common explanation
195 was that Bcd activates *hb* by cooperatively binding multiple sites in the regulatory
196 sequence[33,54,55]. Here, we found that the Hill coefficient of P1-specific GRFs (P1-5'UTR: $h = 7.1$
197 $\pm 0.4$, P1-intron: $h = 7.2 \pm 0.3$, mean ± s.e.m.) was significantly higher than that of P2-specific ($h$

198 = 4.6 ± 0.2) and CDS ($h$ = 5.3 ± 0.3) GRFs during nc11–13 ($p<0.05$, Fig. 3d). Such difference in
199 Hill coefficient is robust against Bcd dosage change (Supplementary Fig. 6c), suggesting that
200 P1 activation corresponds to higher-order cooperative binding than P2 activation. Moreover, P1-
201 specific signals exhibited a higher concentration threshold $C_0$ for activation than P2-specific and
202 CDS signals, with their ratios being consistently larger than one (Fig. 3e). This agrees with our
203 observation of P1 and P2 expression boundaries (Fig. 1g).

204 To directly quantify Bcd binding corresponding to P1 and P2 activation, we measured local
205 enrichment of the Bcd signal in the vicinity of P1- and P2-active *hb* loci[32] (Fig. 3f; see Methods
206 and Supplementary Note). In the anterior expression domain of nc11–12 embryos, we estimated
207 an average binding of ~6.2 Bcd molecules at P1-active loci, exceeding that of ~4.6 Bcd
208 molecules at P2-active loci (Fig. 3f). This result confirmed that P1 activation requires cooperative
209 binding of more Bcd molecules than P2 activation. In nc13, along with the increase in P1 activity
210 (Fig. 1d, f), the number of bound Bcd molecules at P1-active loci dropped to ~5.4. Examining
211 how Bcd binding at active promoter loci varied with nuclear Bcd concentration (or nuclear
212 position) revealed that both P1- and P2-specific binding curves plateaued at ~4–5 Bcd molecules,
213 while the P1-specific binding curve exhibited an additional plateau with ~8–10 Bcd molecules in
214 nc11–12 (Fig. 3g; Supplementary Fig. 7). As suggested previously, these plateaus may
215 correspond to distinct Bcd binding states at *hb* enhancers[32]. The additional Bcd binding plateau
216 for P1 implies that P1 activation may involve more Bcd-binding steps than P2 activation. This
217 additional plateau decreased to ~6 Bcd molecules in nc13 (Supplementary Fig. 7), implying a

218 change in Bcd binding dynamics.

219

220 **Two Bcd-dependent enhancers synergistically drive P1 activation.**

221 Two Bcd-dependent enhancers are involved in early *hb* regulation. To distinguish their roles in
222 P1 and P2 activation, we used transgenic fly lines derived from a bacterial artificial chromosome
223 (BAC) containing the *hb* gene and its regulatory sequence[35] (Fig. 4a; see Methods). The
224 transgenes retained the *hb* promoters, introns, and UTRs, whereas the *hb* CDS was replaced
225 with a *yellow* reporter gene. The distal and proximal enhancers of the two transgenes were
226 substituted, respectively, while the regulatory sequence of the third transgene was kept intact as
227 a control.

228 To measure the promoter activities of the transgene, we labeled transgenic embryos with
229 three sets of smFISH probes targeting P1 intron, *hb* CDS, and *yellow*, respectively (Fig. 4b; see
230 Supplementary Fig. 8 for a complementary study using different probe sets). The P1 signal
231 corresponding to the endogenous *hb* gene was identified and excluded based on its
232 colocalization with the *hb* CDS signal. We compared the anterior P1 and *yellow* (mainly from P2)
233 signals in the control and enhancer-removed transgenes in nc11–13 embryos (Fig. 4c;
234 Supplementary Fig. 9a, b). Removal of either enhancer significantly lowered the percentage of
235 the *yellow*-positive anterior nuclei by ~10% in nc11–12 (Fig. 4d), consistent with previous reports

236 that both enhancers are required for authentic *hb* expression[23,30]. Similarly, removing either
237 enhancer lowered the percentage of P1-active anterior nuclei by >33% in nc11–12 embryos (Fig.
238 4d), suggesting that early P1 activation also relies on both enhancers.

239 Comparing the *yellow* expression profiles between the control and enhancer-removed
240 transgenes (Fig. 4e; Supplementary Fig. 9c, d) revealed a decrease in the maximum anterior
241 expression level $r_{max}$ by ~30-40% in response to proximal or distal enhancer removal in nc11–12
242 (Fig. 4f). This agrees with a previous report that P2 subadditively integrates regulatory inputs
243 from different enhancers via enhancer competition[23]. In contrast, removing the proximal and
244 distal enhancers decreased the maximum expression level $r_{max}$ of P1 by $\geqslant$60% and $\geqslant$37% in
245 nc11–12, respectively (Fig. 4f). It suggests that the two enhancers activate P1 with little
246 competition.

247 Besides affecting the expression amplitude, both enhancers are critical for promoter
248 expression boundaries. Specifically, removing the proximal enhancer shifted the *yellow*
249 expression boundary towards the anterior pole, while removing the distal enhancer caused a
250 posterior shift of the *yellow* profile (Fig. 4g). These results agree with the enhancer competition
251 model for P2 activation[23], in which the expression boundary resulting from two enhancers lies
252 between that from individual ones. In contrast, the deletion of either enhancer caused an anterior
253 shift of the P1 expression boundary (Fig. 4h), which is inconsistent with the enhancer competition
254 model. It shows that the existence of a second enhancer helps activate P1 at lower Bcd

concentrations. Thus, the two enhancers may interact synergistically to drive P1 activation.

Moreover, the Hill coefficient of P1 dropped significantly from ~6.6 to ~4.4 ($p<0.05$), a value close to the Hill coefficient of P2, in response to enhancer removal (Supplementary Fig. 9e). It suggests that synergistic action of the two enhancers involves Bcd binding at both of them. In contrast, the Hill coefficient of P2 does not change significantly with enhancer removal, indicating Bcd binding at a single enhancer.

**Promoter-specific transcription kinetics reveal a unified scheme of enhancer-promoter interaction**

Nascent mRNA copy number statistics reflect the microscopic mechanisms of gene regulation[56]. Previous studies reported a super-Poissonian distribution of the nascent mRNA copy number on individual *hb* gene loci[31,32,45]. Such distribution can be explained by a minimal model of two-state transcription kinetics[42,46,56,57] with Bcd modulating the frequency of stochastic transition from an inactive to an active transcription state[32].

To uncover the kinetics of P1 and P2 transcription, we measured, for each embryo, the distributions of P1 and P2 nascent mRNA signals in different Bcd concentration ranges (Fig. 5a). Contrary to the prediction from the two-state model[46], each promoter exhibited a wide distribution with more than one population of active loci characterized by different expression levels (Fig.

273 5a). A natural explanation of this phenomenon is that some of the observed promoter loci may
274 indeed be a pair of closely located sister loci that are indistinguishable under the microscope[31,45].
275 However, the nascent mRNA signal from individual optically resolved sister loci also exhibited
276 two active populations, with >38% of P1 and >17% of P2 active sister loci corresponding to the
277 minor population (Fig. 5b; see Supplementary Note). Based on the correspondence between
278 distribution modality and the number of transcription states, P1 and P2 transcription should be
279 modeled as a three-state process (see Supplementary Note).

280 In this model, a promoter randomly switches between an inactive state (state 0) and two
281 active states (states 1 and 2) with Poissonian rates $k_{ij}$ ($i$, $j$ = 0, 1, 2), whereas new transcripts are
282 only initiated in active states with rates $k_{\mathrm{INI},i}$. Following transcription initiation, each nascent
283 mRNA elongates with constant speed $V_{\mathrm{EL}}$, and resides on the gene for an extra termination
284 period $T_{\mathrm{R}}$ before being released (Fig. 5a). With a detailed balance between states, we solved
285 the steady-state distribution of nascent mRNA signal per promoter locus and compared it with
286 experimental data to infer the kinetic parameters (Fig. 5a; see Supplementary Note). We found
287 that P1 and P2 both followed a sequential activation scheme from state 0 to state 2
288 (Supplementary Fig. 10a; Supplementary Note). Bcd mainly regulated the activation rates, $k_{01}$
289 and $k_{12}$, while other kinetic rates remained constant (Supplementary Fig. 10b; see
290 Supplementary Fig. 10c, d for a complementary study using the CDS signal as a proxy for P2).

291 For both promoters, the Bcd dependence of $k_{01}$ and $k_{12}$ satisfied Hill functions (Fig. 5c), with

292 the corresponding Hill coefficient $h$ and concentration threshold $C_0$ identical between the
293 promoters (Fig. 5d, e). Specifically, the Hill coefficient of $k_{01}$ was ~4.5 (P1: $h$ = 5.2 ± 0.5, P2: $h$ =
294 4.2 ± 1.0, Fig. 5d), close to that of P2-GRF. In contrast, the Hill coefficient of $k_{12}$ was ~10 (P1: $h$
295 = 10.0 ± 2.8, P2: $h$ = 10.8 ± 3.3, Fig. 5d). $h$ of $k_{01}$ and $k_{12}$ each matched a plateau in the Bcd
296 binding curves of *hb* promoters (Fig. 3g), suggesting that each promoter activation step may
297 involve a cooperative Bcd binding event. The concentration threshold of $k_{01}$ was ~24 nM (P1: $C_0$
298 = 23.5 ± 1.4, P2: $C_0$ = 24.1 ± 2.0, Fig. 5e), which was close to that of P2-GRF. In contrast, the
299 concentration threshold of $k_{12}$ was only ~17 nM (P1: $C_0$ = 17.1 ± 0.2, P2: $C_0$ = 17.2 ± 0.4, Fig.
300 5e), suggesting that the second Bcd-binding event is easier to happen than the first one.

301 In addition to the Bcd dependence, the ratio between the activation and inactivation rates
302 determines the probability of each transcription state. In the anterior expression domain, $k_{01}/k_{10}$
303 for P1 was smaller than that for P2 by ~4–7 folds (Supplementary Fig. 10e), indicating that P2
304 is more likely to be activated than P1. In contrast, $k_{12}/k_{21}$ for P1 exceeded that for P2 by >3 folds
305 (Supplementary Fig. 10f), indicating that active P1 is more inclined to reach state 2 than active
306 P2. For both promoters, the transcription initiation rate of state 2 was larger than that of state 1
307 (Fig. 5f). Therefore, state 2 contributed much more to P1 transcription than to P2 transcription
308 (57.9% vs. 17.5%, Fig. 5g). This explains the difference (in $h$ and $C_0$) between the P1- and P2-
309 GRFs. I.e., the differential regulation of P1 and P2 results from their preference for different
310 active states.

To relate transcription kinetics with enhancer activities, we applied theoretical analysis to enhancer deletion experiments. For early P1 transcription (nc11-12), the removal of either enhancer yielded a uniform decrease of $k_{01}$ by ~30% in the anterior expression domain (Fig. 5h; Supplementary Fig. 10g). Deleting the proximal enhancer also caused a modest anterior shift of the $k_{01}$ boundary (Fig. 5h). These results suggest that either enhancer can drive promoter activation to state 1. In contrast, deleting either enhancer caused a dramatic anterior shift of the $k_{12}$ boundary (Fig. 5h; Supplementary Fig. 10h). This indicates that state 2 is primarily driven by the synergistic action of both enhancers. When the Bcd concentration is sufficiently high, state 2 may also be driven by a single enhancer, possibly due to extra Bcd binding at weak binding sites. These mechanisms explain the difference between P1 and P2 in the expression boundary shift upon enhancer removal (Fig. 4g, h).

Altogether, we propose that the two *hb* promoters follow a common kinetic scheme of Bcd-dependent activation (Fig. 5i; see Supplementary Note). Cooperative Bcd binding at either enhancer can drive a promoter to a basal active state, while full-power transcription requires additional Bcd binding at the second enhancer. The two promoters differ in their responses to different Bcd binding configurations. P2 is primarily activated by Bcd binding at a single enhancer, which results in competitive action of the two enhancers. In contrast, P1 activation relies more on Bcd binding at both enhancers, which leads to synergistic enhancer action. The predicted joint distribution of P1 and P2 activities agreed well with experimental results (Fig. 5j; see Supplementary Note).

## Discussion

Traditional FISH and lacZ reporter experiments reported that Bcd-dependent *hb* transcription in early embryogenesis only involves the P2 promoter[33,34,36,37], while P1 is Bcd-insensitive and remains silent until late nc14[27,30,36]. Here, using smFISH method with single-molecule sensitivity, we showed that Bcd activated both promoters in the anterior domain of nc11–13 embryos. P1 contributes up to ~13% of nascent *hb* signal and up to ~20% of nascent *hb* noise. These fractions are modest compared to those of P2, yet significant enough to affect the expression patterns of *hb*-target genes. Thus, the Bcd-dependent P1 expression is not negligible.

According to previous studies, the difference in P1 and P2 expression levels is due to specific motif codes in promoter sequences. P2 contains Zelda binding sites and a strong TATA box to facilitate chromatin opening and promoter activation[16,30,58]. In contrast, the lack of Zelda binding sites and TATA box in the P1 promoter significantly impedes local chromatin opening and promoter activation[17,38,59]. However, our results suggest that the lack of these motif codes may be insufficient to completely block early P1 activity. In fact, the local chromatin state of a gene has been reported to be highly dynamic and can randomly switch between “open” and “close” to allow transient gene expression[13,60]. This effect has been proposed as a mechanism for

349 transcriptional bursting[60-62], consistent with our observation of bursty P1 transcription. Moreover,
350 our observation that the percentage of active P1 loci increased with the nuclear cycle suggests
351 a gradual increase in chromatin opening frequency during development. The stronger P1 activity
352 in late nc14 may be a continuation of this trend.

353 In addition to the expression amplitude, the Bcd dependence of P1 and P2 showed
354 quantitative differences. P1 activation requires a higher Bcd concentration and binding of more
355 Bcd molecules than P2 activation. One possible mechanism for this difference is that the two
356 promoters are driven by different enhancers. However, we showed that early P1 and P2
357 activation relied on the same pair of enhancers[35]. Another possibility is that the two promoters
358 distinguish their behaviors by mutual interaction[63]. However, we found little correlation between
359 P1 and P2 activities. This is consistent with a previous promoter deletion experiment[30] and
360 indicates no interaction between the two promoters. Thus, the difference in Bcd dependence
361 originates from complex enhancer-promoter interactions.

362 One form of complex enhancer-promoter interaction is the competitive action of the proximal
363 and shadow enhancers for *hb* P2 activation[23]. Here, each enhancer corresponds to a unique
364 Bcd-dependent regulatory effect, and the two enhancers compete for promoter activation[12]. The
365 resulting P2 expression is a sub-additive integration of individual enhancers' regulatory effect.
366 This mechanism can lead to a different Bcd-dependent relationship if the enhancers' relative
367 weights in the integration change. However, the P1 expression boundary always shifts anteriorly

368 in response to the deletion of either enhancer, suggesting a synergistic, as opposed to
369 competitive, action of the two enhancers for P1 activation. This result also revealed that the Bcd
370 dependence of P1 and P2 differed even in the case of a single enhancer, indicating that the TF
371 dependence of an enhancer's regulatory effect may be promoter-specific.

372 To understand the mechanism of promoter-specific interactions with enhancers, we analyzed
373 the transcription kinetics of individual promoters. Unlike the previous model of two-state *hb*
374 transcription[31,32,45], we found that both P1 and P2 activities need to be described by three-state
375 kinetics with sequential activation steps. Multi-step transcriptional activation was previously
376 proposed to model enhancer-, promoter-, or chromatin-related intermediate states during
377 activation[13,60,62,64,65]. However, these intermediate states are rarely experimentally detectable. In
378 the case of *hb*, previous measurements mixed signals from P1 and P2, whose difference easily
379 overwhelmed the subtle signatures of different transcription states. Here, the identification of
380 multiple active states was possible owing to promoter-specific demixing of the *hb* nascent mRNA
381 signal. Thus, it is important to distinguish alternative promoters in the study of complex gene
382 regulation.

383 Our model revealed that Bcd regulates both activation steps ($k_{01}$ and $k_{12}$) through cooperative
384 binding. This suggests that the two active states of a promoter may correspond to different Bcd
385 binding states at *hb*. Previous literature has shown evidence of multiple Bcd binding states at
386 *hb*[32]. However, only one binding state (with ~4–5 Bcd molecules) was reported to coincide with

387 gene activation[32]. The biological functions of the other binding states were unclear. Here, we
388 showed that the binding of ~4–5 Bcd molecules happened at a single enhancer (either proximal
389 or distal). Enhancers at this binding state can competitively drive promoter activation, consistent
390 with a previous study of *hb* P2 activation[23]. In contrast, the binding of ~8–10 Bcd molecules
391 involved both enhancers, leading to synergistic activation of a promoter. Such a mechanism was
392 proposed for some eukaryotic genes[23,66] but has never been reported for *hb*.

393 P1 and P2 differ in their preference for different Bcd binding states. Specifically, the P1
394 response to the lower Bcd binding state is much less than that of P2. A possible reason is that
395 P1 lacks specific motif codes for chromatin opening[30]. Thus, its activation may require TF binding
396 at enhancers to help open the local chromatin configuration[61,62]. Binding of more Bcd molecules
397 at both enhancers may be more effective for chromatin opening. However, since P1 locates
398 between the two enhancers, binding of Bcd (and other regulatory factors) at the proximal
399 enhancer may physically block P1 transcription. It is unclear how P1 coordinates its activation
400 and transcription. One possibility is that TF binding at enhancers is only transiently needed to
401 form the preinitiation complex. Future experiments using high-resolution live imaging techniques
402 would likely solve this puzzle[67,68].

403 Altogether, these results showed that a single kinetic scheme could create apparently
404 different types of enhancer-promoter interactions. Such a unified scheme may be shared by
405 alternative promoter regulation in many eukaryotic systems and may be crucial for phenotypic

complexity in higher eukaryotes[7]. Moreover, our combined experimental and theoretical approach directly relates TF binding at individual enhancers with the stochastic activity of each promoter. A generalization of this approach (e.g., by including more regulatory factors) will help understand the regulatory mechanisms for cell fate decisions and enable the precise design of synthetic gene circuits[69].

## Methods

### Fly strains

Oregon-R (OreR) strain was used as the wild type. CRISPR mutant strains with P1 and P2 deletions (ΔP1C and ΔP2C) were developed previously[30] and were obtained as gifts from Dr. Stephen Small (New York University) and Dr. Pinar Onal (Northwestern University). 1×*bcd* strain (*+/CyO-bcd+*; *E1s*) was developed previously[51] and was obtained as a gift from Dr. Jun Ma (Zhejiang University). *hb*-BAC reporter constructs were previously developed[35]. The distal-enhancer-removed BAC construct and its control were obtained as gifts from Dr. Michael Perry (University of California San Diego) and Dr. Alistair Boettiger (Stanford University). The proximal-enhancer-removed construct and its control were rebuilt as previously described[35]. Transgenes in these constructs were constructed from BAC CH322-55J23, which contains a 20-kbp *Drosophila* genomic sequence encompassing the *hb* gene and its proximal and distal enhancers.

425 A *yellow-kanamycin* fusion was used to replace the *hb* CDS in all BACs, while the proximal or
426 distal enhancers in the BACs were substituted with *ampicillin*. All BACs were integrated on
427 chromosome 2 of *Drosophila*. The distal-enhancer-removed transgene and its control were
428 integrated into landing site VK37 (Bloomington Stock Center number 24872). The proximal-
429 enhancer-removed transgene and its control were integrated into landing site attP40
430 (Bloomington Stock Center number 25709).
431

432 **smFISH probe design**

433 Sets of DNA oligonucleotides complementary to the target transcripts (nine probes for *hb* P1
434 5'UTR, 32 probes for *hb* P1 intron, 48 probes for *hb* CDS, eight probes for *hb* P2 3'UTR, 43
435 probes for *yellow*, 33 probes for *Kr*, and 29 probes for *kni*) were designed (Supplementary Table
436 1) and synthesized (Biosearch Technologies). Each probe was ordered with a 3' amine group
437 (mdC(TEG-Amino)) and was conjugated to various fluorophores, as described previously[32]. In
438 most experiments, *hb* P1-5'UTR and intron probes were conjugated with tetramethylrhodamine
439 (TAMRA; Thermo Fisher Scientific, C6123). *hb* P2-3'UTR, *yellow,* and *Kr* probes were
440 conjugated with Alexa Fluor™ 647 (Invitrogen, A20106). *kni* probes were conjugated with Alexa
441 Fluor™ 488 (Invitrogen, A20100). *hb* CDS probes were conjugated with either Alexa Fluor™ 647
442 (Invitrogen, A20106) or Alexa Fluor™ 488 (Invitrogen, A20100). In a control experiment, P1-
443 5'UTR probes with Alexa Fluor™ 647 and P2-3'UTR probes with TAMRA were used to label WT
444 embryos. A comparison between signals from the same probe set with different fluorophores

445 revealed that the efficiencies of different fluorescent detectors were comparable (Supplementary
446 Fig. 2e, f).
447
448 **Embryo staining**
449 Embryo collection, fixation, and labeling were performed according to a previously published
450 protocol[32]. Briefly, 2-hour-old embryos were collected at 25°C, fixed with 4% paraformaldehyde
451 solution, and stored in 100% methanol at −20°C. For smFISH, fixed embryos were rehydrated
452 (4 × 10 min) in PBTx (1× PBS, 0.1% (v/v) Triton X-100) at room temperature, washed (2 × 10
453 min) in hybridization wash buffer at 30°C, and incubated with the probe-containing hybridization
454 buffer (2× SSC, 20% (w/v) formamide, 0.1% (v/v) Triton X-100) at 30°C overnight. After
455 hybridization, embryos were washed in hybridization wash buffer at 30°C (2 × 10 min) and in 2×
456 SSC at room temperature (2 × 10 min). For immunofluorescence (IF), embryos were washed (4
457 × 10 min) in PBTx and blocked in PBT-B (1× PBS, 20% (v/v) western blocking reagent (Roche,
458 11921673001), 2 mM ribonucleoside vanadyl complex (NEB, S1402S), 0.1% (v/v) Triton X-100)
459 at room temperature for 1 h. The preabsorbed rabbit anti-Bcd primary antibody (Santa Cruz
460 Biotechnology, SC-66818) was incubated with embryos at 4°C for 20 h. Following primary
461 antibody staining, embryos were further washed (4 × 10 min) in PBTx, blocked in PBT-B at room
462 temperature for 1 h, and incubated with goat anti-rabbit IgG secondary antibodies conjugated
463 with Alexa Fluor™ 488 (Invitrogen, A11008) at room temperature for 1 h. For DNA
464 counterstaining, embryos were washed (4 × 10 min) in PBTx and stained with Hoechst 33342 at

465 room temperature for 10 min. Following additional washes (4 × 10 min) in PBTx, embryos were
466 mounted in Aqua-Poly/Mount (Polysciences, 18606). Imaging was performed after the samples
467 were completely solidified.
468

469 **HCR-FISH probe design and labeling**

470 HCR-FISH was used as a replacement of smFISH to identify active *yellow* or *lacZ* loci in some
471 control experiments (Supplementary Fig. 4, 9). Design of HCR-FISH probes and labeling of
472 embryos are described in detail in Supplementary Note.
473

474 **Imaging**

475 Most embryos were imaged using a Zeiss LSM 880 laser scanning confocal microscope
476 equipped with a GaAsP detector and a 63× oil-immersion objective (1.4 NA). 16-bit image stacks
477 were acquired with a pixel size of 71 × 71 $nm^2$ and a z-step size of 0.32 μm. A small number of
478 embryos (n = 5) were imaged using a Leica TCS SP8 confocal microscope equipped with a
479 GaAsP detector and a 63× oil-immersion objective (1.4 NA). 12-bit image stacks were acquired
480 with a pixel size of 81 × 81 $nm^2$ and a z-step size of 0.3 μm. To image *hb* transcription, nc11–13
481 embryos at the mitotic interphase were selected based on the number and shape of the nuclei
482 (Hoechst signal). Pre-gastrulation nc14 embryos were selected to image *Kr* and *kni* activities.
483 Approximately 10 μm of the cortex layer of each embryo was imaged. For promoter mutant lines,
484 each mutant allele was made heterozygous with a TM3 Sb balancer containing a *hb-lacZ*

485 reporter. To identify homozygous mutant embryos, the *lacZ* FISH signal (labeled by Alexa Fluor™
486 594) was imaged first, and those embryos without active *lacZ* loci were selected for subsequent
487 imaging.
488

489 **Preprocessing and nuclear segmentation**

490 Image processing and data analysis followed a previously developed pipeline[32] with updated
491 algorithms to improve accuracy and efficiency. Briefly, raw images were divided by a normalized
492 flat-field image to correct for monochromatic aberrations. Three-dimensional (3D) segmentation
493 of nuclei from the Hoechst image stack was done using a combination of local threshold (to
494 optimize a circularity parameter) and watershed (to separate the merged nuclei). The nuclear
495 cleavage cycle of the embryo was determined by the number of recognized nuclei. Embryo
496 boundary was identified from the averaged Hoechst image by thresholding image pixels outside
497 the nuclear area. This boundary was then used to determine the AP position of each nucleus.
498

499 **mRNA quantification**

500 Spot candidates in smFISH images were identified as 3D local maxima in the image stack. Since
501 the splicing process occurs inside the nucleus, intron spot candidates were only identified in the
502 nuclear region. The local intensity profile of each candidate was fitted to a two-dimensional (2D)
503 Gaussian function to extract the peak height ($I_{peak}$) and radius ($\sigma_0$). The spot intensity was
504 calculated as $I = 2\pi I_{peak}\sigma_0^2$. By comparing the joint distribution of peak height and radius between

505 the anterior and posterior spots, a 2D threshold was determined to distinguish real mRNA spots
506 from background noise (Supplementary Fig. 2a). The typical intensity, $I_0$, of a single mRNA
507 molecule was extracted by fitting the primary peak of the spot intensity distribution to a multi-
508 Gaussian function (Supplementary Fig. 2b). A threshold of $3I_0$ was defined to identify sites of
509 active transcription from mRNA spots inside each nucleus. The equivalent number of nascent
510 transcripts at each transcription site was estimated by dividing the intensity of the transcription
511 site by $I_0$.

512 To identify signals corresponding to the same *hb* locus in two different smFISH channels, we
513 calculated the mutual distance between every possible pair of active transcription sites detected
514 in different smFISH channels. The distribution of all the mutual distances in an embryo exhibited
515 a clear population within a small distance range (<0.55 μm), outside which the distribution
516 became flat (Supplementary Fig. 2c). This population corresponded to colocalized pairs of active
517 transcription sites. Using a threshold distance of 0.55 μm, we identified these colocalized pairs,
518 each of which belonged to a single *hb* locus. In contrast, each unpaired active transcription site
519 belonged to a different *hb* locus, whose signal in the other smFISH channel was zero.

520 In some anterior nuclei, replication of the *hb* gene can lead to >2 bright FISH spots[31]. To
521 identify nascent mRNA signals corresponding to newly replicated sister loci pairs, we calculated
522 the mutual distance between every possible pair of active transcription sites in each nucleus.
523 The distribution of all the mutual distances in the embryo exhibited two distinct populations

524 (Supplementary Fig. 2d). The population with smaller mutual distances corresponds to sister loci
525 pairs, while the other population corresponds to unpaired homologous loci. Using a threshold
526 distance of 0.71 μm that lay at the valley point between the two populations in the distribution,
527 sister loci pairs were distinguished from unpaired homologous loci.

528 Nuclei with and without paired active loci were both analyzed in the paper. For single-locus
529 analysis, we roughly estimated the number of silent loci in each nucleus using a default criterion.
530 I.e., nuclei with paired active loci are post-replication and have four copies of the *hb* gene, while
531 nuclei without paired active loci were coarsely assumed to be pre-replication and have two
532 copies of the *hb* gene. To avoid the possible inaccuracy in gene copy inference, we summed
533 over the nascent signal in each nucleus as an alternative measure of promoter activity for mean-
534 level analysis.

535

536 **Protein quantification**

537 The average immunofluorescence (IF) intensity of each nucleus was calculated from the central
538 *z*-slice of the nucleus. IF spots in the cytoplasm were identified and quantified to determine the
539 typical intensity, $I_1$, of a single protein molecule, following the same procedure used for the
540 smFISH signal (Supplementary Fig. 6a). The absolute protein concentration of a nucleus was
541 estimated by dividing the average IF intensity of the nucleus by $(2\pi)^{1/2}\sigma_z I_1$, where $\sigma_z$ is the half-
542 width of the single-protein intensity profile in the *z* dimension.

In the preceding steps, nuclear segmentation, embryo boundary identification, and detection of active transcription sites could be further refined and corrected manually using custom MATLAB graphical user interfaces.

**Measuring the spatial profile of promoter activity**

We analyzed the expression profile of a promoter using embryos in mid-to-late mitotic interphase to ensure steady-state promoter activity. For each embryo, we plotted the nascent mRNA signal ($r$) against nuclear position ($x$) for all nuclei and binned individual data points by $x$. Within the range of 0.25–0.75 EL, we used a least-squares algorithm to fit the binned data to a logistic function:

$$r = r_{\max} \frac{e^{-(x-x_0)/d}}{e^{-(x-x_0)/d} + 1} + r_0 \quad (3)$$

where $x_0$ is the boundary position of the expression domain, $d$ is the half-width of the transition region, $r_{\max}$ is the maximal transcription level induced in the anterior region, and $r_0$ denotes the basal activity in the posterior part.

**Measuring the fluctuation of promoter activity**

At the single-molecule level, gene expression constantly varies over time and between different cells. Based on the correlation between the homologous loci in the same cell, the fluctuation or

561 noise of gene expression may be divided into two parts: the intrinsic noise due to the inherent
562 stochasticity of biochemical reactions and the extrinsic noise caused by cell-to-cell variability of
563 the microenvironment[70].

564 To characterize the expression variability of different *hb* promoters, we computed, for each
565 promoter, the Fano factor ($F$) of nascent mRNA signal ($r$, in units of the number of molecules)
566 per locus in the anterior expression domain (0.2–0.4 EL)[32]:

567 $$F = \frac{\sigma^2}{\langle r \rangle} \tag{4}$$

568 where $\langle r \rangle$ and $\sigma$ are the mean and standard deviation of the single-locus data, respectively. The
569 Fano factors of P1- and P2-specific signals were much larger than one (Fig. 1h), indicating bursty
570 transcription from both promoters[42,43].

571 We quantified the intrinsic noise of P1 and P2 expression in the anterior expression domain
572 (0.2–0.4 EL) of an embryo using the following formula[44]:

573 $$\eta^2 = \frac{\langle (m_1 - m_2)^2 \rangle}{2\langle m_1 \rangle \langle m_2 \rangle} \tag{5}$$

574 where $m_1$ and $m_2$ are nascent mRNA signals at two homologous loci in the same nucleus
575 measured using a given probe set, respectively.

576

577 **Measuring the correlation between promoter loci**

578 To distinguish the intrinsic and extrinsic noise of different promoter activities, we computed
579 the correlation coefficient ($\rho$) of the nascent mRNA signal between the two homologous copies
580 of a given promoter within the same nucleus. Specifically, we divided the single-locus data of the
581 promoter activity (in units of the number of mRNA molecules) in an embryo into two groups, $r_1$
582 and $r_2$. Each group corresponded to one of the two homologous loci in the nucleus. In a given
583 region of the embryo, we applied the following formula:

584 $$\rho = \frac{\left\langle \left(r_1 - \langle r_1 \rangle\right) \cdot \left(r_2 - \langle r_2 \rangle\right) \right\rangle}{\sigma_1 \sigma_2} \quad (6)$$

585 where $\langle r_1 \rangle$, $\langle r_2 \rangle$, $\sigma_1$, and $\sigma_2$ are the mean and standard deviation of each group, respectively. In
586 the anterior expression domain (0.2–0.4 EL), P1, P2, and CDS signals showed little correlation
587 (<0.15; Supplementary Fig. 3c), agreeing with previous reports of loci independence[31,32]. Thus,
588 intrinsic noise dominates P1 and P2 expression. To further evaluate the interaction between the
589 two promoters, we computed the correlation coefficient between P1 and P2 signals from the
590 same (intra-allele) or different (inter-allele) *hb* loci in the nucleus using the same formula, where
591 $r_1$ and $r_2$ denote the activities of individual P1 and P2 copies, respectively. Both quantities were
592 at low levels (<0.06; Supplementary Fig. 3d), indicating that the two promoters do not interact
593 during expression.

594

**Estimating promoter contributions to *hb* activity**

To estimate the contributions of P1 and P2 activities to nascent *hb* transcription, we co-labeled fly embryos with P1-5'UTR, P2-3'UTR, and CDS probes. For each nucleus in the position range of 0.3–0.6 EL of an embryo, we computed the nascent signal of each probe set (in units of the number of mRNA molecules). By locally averaging the signal over the nearest five neighboring nuclei, we plotted the three FISH signals against each other and fitted the data to Equation (1) using linear regression. $a_1$ extracted from most embryos at mitotic interphase exhibited a typical value of ~0.5 (Supplementary Fig. 5a), close to the ratio between CDS and P1-5'UTR signals in ΔP2C mutant (Supplementary Fig. 4a, b; see Supplementary Note). In contrast, embryos close to mitosis gave $a_1 \gtrsim 1$, as transcription initiation has been turned off at that moment (Supplementary Fig. 5a). These embryos were excluded from analysis using a threshold $a_1$ of 0.8.

With the inferred ratio parameters $a_1$ and $a_2$, we scaled nuclear P1-5'UTR and P2-3'UTR signals to compare with nuclear CDS signal in different parts of the embryo (Supplementary Fig. 5c). The two scaled signals reconstructed most of the CDS signal in the anterior expression domain (0.2–0.4 EL). We computed the average P1 contribution to the anterior nascent CDS signal as

$$f_{\text{P1-CDS}} = \frac{a_1 r_{\text{P1}}}{r_{\text{CDS}}} \tag{7}$$

where $r_{\text{CDS}}$ and $r_{\text{P1}}$ are the average nascent CDS and P1-5'UTR signals per nucleus, respectively.

The rest of the nascent CDS signal came from P2. Extra CDS component existed in terminal regions (0–0.2 and 0.8–1 EL; Supplementary Fig. 5c), suggesting different ratio parameters for these regions.

To compute promoter contributions to the intrinsic noise of nascent *hb* transcription, we estimated $\eta^2_{\text{P1-CDS}}$ in Equation (3) from the intrinsic noise of the nascent P1-5'UTR signal (see Supplementary Note). The contribution of P1 to the intrinsic noise of nascent CDS signal was calculated as:

$$w_{\text{P1-CDS}} = \frac{\eta^2_{\text{P1-CDS}} f^2_{\text{P1-CDS}}}{\eta^2_{\text{CDS}}} \tag{8}$$

The rest of the intrinsic noise was contributed by P2.

**Measuring the expression patterns of *Kr* and *kni***

We measured the expression patterns of *Kr* and *kni* using pre-gastrulation nc14 embryos. For each embryo, we plotted the nascent mRNA signal ($r$, in units of the number of molecules) against nuclear position ($x$) for all nuclei and binned individual data points by $x$. Within position ranges 0.3–0.8 EL (for *Kr*) or 0.4–1.0 EL (for *kni*), we used a least-squares algorithm to fit the binned data to a product of two logistic functions:

$$r = r_{\max} \frac{e^{(x-x_{0\text{A}})/d_{\text{A}}}}{(e^{(x-x_{0\text{A}})/d_{\text{A}}}+1)} \frac{e^{-(x-x_{0\text{P}})/d_{\text{P}}}}{(e^{-(x-x_{0\text{P}})/d_{\text{P}}}+1)} + r_0 \tag{9}$$

where $x_{0A}$ and $x_{0P}$ denote the anterior and posterior boundary positions of the expression domain, respectively.

**Measuring the gene regulation function (GRF)**

To analyze the regulation of a promoter by Bcd, we plotted, for each embryo, the nascent mRNA signal versus Bcd concentration for each nucleus within the position range of 0.25–0.75 EL. To extract the GRF, we binned individual data points by Bcd concentration and fitted them to a Hill function using a least-squares algorithm:

$$y = a\frac{[\mathrm{Bcd}]^h}{[\mathrm{Bcd}]^h + C_0^h} + d \tag{10}$$

where $h$ is the Hill coefficient, $C_0$ is the concentration threshold for promoter activation, $a$ indicates the maximal level of Bcd-dependent activity, and $d$ denotes the basal activity.

**Measuring Bcd binding**

To quantify Bcd binding at a specific *hb* promoter, we used active P1 and P2 transcription sites to locate individual (active) promoter loci. Near each locus, a “locus-integration region” (*xy* distance ≤ 3 pixel and *z* distance = 0 from the locus) and an “out-of-locus region” (*xy* distance ≤ 6 pixels and ≥ 3 pixels, *z* distance = 0) were defined, which covered the nuclear volumes of $V_I$ and $V_o$, respectively. The nuclear IF signal (in units of the number of Bcd molecules) within these

649 two regions was integrated and denoted as $I_l$ and $I_o$, respectively. The enriched Bcd signal was
650 defined as the difference between $I_l$ and $I_o$ in consideration of the volume difference between the
651 two regions, i.e., $I_{enrich} = I_l - I_o \cdot V_l / V_o$. Data from multiple embryos in the sample nuclear cycle
652 were pooled to increase the sample size. Data in the nuclear position range of 0.2–0.35EL were
653 averaged to estimate the mean Bcd binding level of the anterior expression domain. To plot the
654 Bcd binding curve, we binned the single-locus data by nuclear position and related mean Bcd
655 enrichment with mean nuclear position or Bcd concentration.
656

657 **Mathematical modeling of transcriptional kinetics**

658 Stochastic modeling and inference of transcriptional kinetics are described in detail in
659 Supplementary Note.
660

661 **Data and code availability**

662 The raw image data reported in this paper are available at http://gofile.me/4yuzx/wKna2V9pK.
663 Source data are provided with this paper. Custom scripts for data analysis and mathematical
664 modeling were written in MATLAB 2018a (MathWorks) and are available at
665 https://github.com/Dr-xu-lab/Quantify-the-transcriptional-regulation.
666

667 **References**

668 1. Levine, M. Transcriptional enhancers in animal development and evolution. *Curr. Biol.* **20**, 754-763
669 (2010).

670
671 2. Vernimmen, D. & Bickmore, W. A. The Hierarchy of Transcriptional Activation: From Enhancer to
672 Promoter. *Trends Genet.* **31**, 696-708 (2015).
673
674 3. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation.
675 *Nat. Rev. Mol. Cell Biol.* **19**, 621-637 (2018).
676
677 4. Rach, E. A., Yuan, H. Y., Majoros, W. H., Tomancak, P. & Ohler, U. Motif composition, conservation
678 and condition-specificity of single and alternative transcription start sites in the Drosophila genome.
679 *Genome Biol.* **10**, R73 (2009).
680
681 5. Carninci, P.*, et al.* Genome-wide analysis of mammalian promoter architecture and evolution. *Nat.*
682 *Genet.* **38**, 626-635 (2006).
683
684 6. Davuluri, R. V., Suzuki, Y., Sugano, S., Plass, C. & Huang, T. H. The functional consequences of
685 alternative promoter use in mammalian genomes. *Trends Genet.* **24**, 167-177 (2008).
686
687 7. Landry, J. R., Mager, D. L. & Wilhelm, B. T. Complex controls: the role of alternative promoters in
688 mammalian genomes. *Trends Genet.* **19**, 640-648 (2003).
689
690 8. Pozner, A.*, et al.* Developmentally regulated promoter-switch transcriptionally controls Runx1
691 function during embryonic hematopoiesis. *BMC Dev. Biol.* **7**, 84 (2007).
692
693 9. Lu, D., Sin, H. S., Lu, C. & Fuller, M. T. Developmental regulation of cell type-specific transcription by
694 novel promoter-proximal sequence elements. *Genes Dev.* **34**, 663-677 (2020).
695
696 10. Sendoel, A.*, et al.* Translation from unconventional 5' start sites drives tumour initiation. *Nature* **541**,
697 494-499 (2017).
698
699 11. de Klerk, E. & t Hoen, P. A. Alternative mRNA transcription, processing, and translation: insights from
700 RNA sequencing. *Trends Genet.* **31**, 128-139 (2015).
701
702 12. Kvon, E. Z., Waymack, R., Gad, M. & Wunderlich, Z. Enhancer redundancy in development and disease.
703 *Nat. Rev. Genet.* **22**, 324-336 (2021).
704
705 13. Voss, T. C. & Hager, G. L. Dynamic regulation of transcriptional states by chromatin and transcription
706 factors. *Nat. Rev. Genet.* **15**, 69-81 (2014).
707
708 14. Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nat.*
709 *Rev. Genet.* **13**, 613-626 (2012).
710

15. Hnisz, D., Day, D. S. & Young, R. A. Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell* **167**, 1188-1200 (2016).

16. Foo, S. M., *et al.* Zelda potentiates morphogen activity by increasing chromatin accessibility. *Curr. Biol.* **24**, 1341-1346 (2014).

17. Sun, Y., *et al.* Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res.* **25**, 1703-1714 (2015).

18. Li, X. Y., Harrison, M. M., Villalta, J. E., Kaplan, T. & Eisen, M. B. Establishment of regions of genomic activity during the Drosophila maternal to zygotic transition. *Elife* **3**, e03737 (2014).

19. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341-1345 (2018).

20. Ghavi-Helm, Y., *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96-100 (2014).

21. Oudelaar, A. M., *et al.* Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* **50**, 1744-1751 (2018).

22. Scholes, C., Biette, K. M., Harden, T. T. & DePace, A. H. Signal Integration by Shadow Enhancers and Enhancer Duplications Varies across the Drosophila Embryo. *Cell Rep.* **26**, 2407-2418 (2019).

23. Bothma, J. P., *et al.* Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. *Elife* **4**, e07956 (2015).

24. Fukaya, T., Lim, B. & Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358-368 (2016).

25. Lim, B., Heist, T., Levine, M. & Fukaya, T. Visualization of Transvection in Living Drosophila Embryos. *Mol. Cell* **70**, 287-296 (2018).

26. Su, W., Jackson, S., Tjian, R. & Echols, H. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes Dev.* **5**, 820-826 (1991).

27. Schroder, C., Tautz, D., Seifert, E. & Jackle, H. Differential regulation of the two transcripts from the Drosophila gap segmentation gene *hunchback*. *EMBO J.* **7**, 2881-2887 (1988).

28. Perry, M. W., Bothma, J. P., Luu, R. D. & Levine, M. Precision of *hunchback* expression in the Drosophila embryo. *Curr. Biol.* **22**, 2247-2252 (2012).

29. Driever, W. & Nusslein-Volhard, C. The bicoid protein is a positive regulator of *hunchback* transcription in the early Drosophila embryo. *Nature* **337**, 138-143 (1989).

30. Ling, J., Umezawa, K. Y., Scott, T. & Small, S. Bicoid-Dependent Activation of the Target Gene *hunchback* Requires a Two-Motif Sequence Code in a Specific Basal Promoter. *Mol. Cell* **75**, 1178-1187 (2019).

31. Little, S. C., Tikhonov, M. & Gregor, T. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789-800 (2013).

32. Xu, H., Sepúlveda, L. A., Figard, L., Sokac, A. M. & Golding, I. Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat. Methods* **12**, 739–742 (2015).

33. Driever, W., Thoma, G. & Nusslein-Volhard, C. Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. *Nature* **340**, 363-367 (1989).

34. Struhl, G., Struhl, K. & Macdonald, P. M. The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57**, 1259-1273 (1989).

35. Perry, M. W., Boettiger, A. N. & Levine, M. Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. *Proc. Natl. Acad. Sci. U S A* **108**, 13570-13575 (2011).

36. Margolis, J. S.*, et al.* Posterior stripe expression of *hunchback* is driven from two promoters by a common enhancer element. *Development* **121**, 3067-3077 (1995).

37. Wu, X., Vasisht, V., Kosman, D., Reinitz, J. & Small, S. Thoracic patterning by the Drosophila gap gene *hunchback*. *Dev. Biol.* **237**, 79-92 (2001).

38. Blythe, S. A. & Wieschaus, E. F. Establishment and maintenance of heritable chromatin structure during early Drosophila embryogenesis. *Elife* **5**, e20148 (2016).

39. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of single RNA transcripts in situ. *Science* **280**, 585-590 (1998).

40. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877-879 (2008).

41. Tautz, D.*, et al.* Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of Drosophila segmentation genes. *Nature* **327**, 383-389 (1987).

42. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**, e309 (2006).

43. Sanchez, A. & Golding, I. Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188-1193 (2013).

44. Waymack, R., Fletcher, A., Enciso, G. & Wunderlich, Z. Shadow enhancers can suppress input transcription factor noise through distinct regulatory logic. *Elife* **9**, e59351 (2020).

45. Zoller, B., Little, S. C. & Gregor, T. Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting. *Cell* **175**, 835-847 (2018).

46. Xu, H., Skinner, S. O., Sokac, A. M. & Golding, I. Stochastic Kinetics of Nascent RNA. *Phys. Rev. Lett.* **117**, 128101 (2016).

47. Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **15**, 163-175 (2014).

48. Lenstra, T. L., Rodriguez, J., Chen, H. & Larson, D. R. Transcription Dynamics in Living Cells. *Annu. Rev. Biophys.* **45**, 25-47 (2016).

49. Hulskamp, M., Pfeifle, C. & Tautz, D. A morphogenetic gradient of *hunchback* protein organizes the expression of the gap genes Kruppel and knirps in the early Drosophila embryo. *Nature* **346**, 577-580 (1990).

50. Struhl, G., Johnston, P. & Lawrence, P. A. Control of Drosophila body pattern by the *hunchback* morphogen gradient. *Cell* **69**, 237-249 (1992).

51. Liu, J. & Ma, J. Dampened regulates the activating potency of Bicoid and the embryonic patterning outcome in Drosophila. *Nat. Commun.* **4**, 2968 (2013).

52. Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Gene regulation at the single-cell level. *Science* **307**, 1962-1965 (2005).

53. Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the limits to positional information. *Cell* **130**, 153-164 (2007).

54. Ma, X., Yuan, D., Diepold, K., Scarborough, T. & Ma, J. The Drosophila morphogenetic protein Bicoid binds DNA cooperatively. *Development* **122**, 1195-1206 (1996).

55. Estrada, J., Wong, F., DePace, A. & Gunawardena, J. Information Integration and Energy Expenditure

in Gene Regulation. *Cell* **166**, 234-244 (2016).

56. Zenklusen, D., Larson, D. R. & Singer, R. H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **15**, 1263-1271 (2008).

57. Peccoud, J. & Ycart, B. Markovian Modeling of Gene-Product Synthesis. *Theor. Popul. Biol.* **48**, 222-234 (1995).

58. Schulz, K. N.*, et al.* Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. *Genome Res.* **25**, 1715-1726 (2015).

59. Nien, C. Y.*, et al.* Temporal coordination of gene networks by Zelda in the early Drosophila embryo. *PLoS Genet.* **7**, e1002339 (2011).

60. Bintu, L.*, et al.* Dynamics of epigenetic regulation at the single-cell level. *Science* **351**, 720-724 (2016).

61. Lammers, N. C.*, et al.* Multimodal transcriptional control of pattern formation in embryonic development. *Proc. Natl. Acad. Sci. U S A* **117**, 836-847 (2020).

62. Eck, E.*, et al.* Quantitative dissection of transcription in development yields evidence for transcription-factor-driven chromatin accessibility. *Elife* **9**, e56429 (2020).

63. Li, X.*, et al.* Heritable, Allele-Specific Chromosomal Looping between Tandem Promoters Specifies Promoter Usage of SHC1. *Mol. Cell. Biol.* **38**, e0065817 (2018).

64. Pimmett, V. L.*, et al.* Quantitative imaging of transcription in living Drosophila embryos reveals the impact of core promoter motifs on promoter state dynamics. *Nat. Commun.* **12**, 4504 (2021).

65. Tantale, K.*, et al.* A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* **7**, 12248 (2016).

66. Choi, J.*, et al.* Evidence for additive and synergistic action of mammalian enhancers during cell fate determination. *Elife* **10**, e65381 (2021).

67. Li, J. R.*, et al.* Single-Molecule Nanoscopy Elucidates RNA Polymerase II Transcription at Single Genes in Live Cells. *Cell* **178**, 491-506 (2019).

68. Li, J.*, et al.* Single-gene imaging links genome topology, promoter-enhancer communication and transcription control. *Nat. Struct. Mol. Biol.* **27**, 1032-1040 (2020).

69. Doshi, J., Willis, K., Madurga, A., Stelzer, C. & Benenson, Y. Multiple Alternative Promoters and

875 Alternative Splicing Enable Universal Transcription-Based Logic Computation in Mammalian Cells.
876 *Cell Rep.* **33**, 108437 (2020).
877
878 70. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell.
879 *Science* **297**, 1183-1186 (2002).
880

881

882 **Acknowledgments**

883 We thank Stephen Small, Pinar Onal, Jun Ma, Michael Perry, and Alistair Boettiger for the
884 generous gift of fly lines. We thank Ido Golding, Anna Sokac, and Jun Ma for insightful discussion
885 and valuable comments on the manuscript. This work was supported by the National Key R&D
886 Program of China (grant no. 2018YFC0310803), the National Science Foundation of China
887 (grant no. 11774225, 41921006), the Natural Science Foundation of Shanghai (grant no.
888 18ZR1419800), and the Burroughs Wellcome Fund Career Award at the Scientific Interface
889 (grant no. 1013907). We gratefully acknowledge the imaging and computing resources provided
890 by the Instrumental Analysis Center and the Student Innovation Center at Shanghai Jiao Tong
891 University.

892

893

894 **Author contributions**

895 Conceptualization, J.W. and H.X.; Methodology, J.W., H.L., S.Z., and H.X.; Software, S.Z. and
896 H.X.; Formal Analysis, J.W. and S.Z.; Investigation, J.W., S.Z., and H.X.; Writing – Original Draft,
897 J.W. and H.X.; Writing – Revised Draft, J.W. and H.X.; Funding Acquisition, H.X.; Resources,
898 H.X.; Supervision, H.X.

899

900

901 **Competing interests**

902 The authors declare no competing interests.

903

904 **Figure legends**

905 **Fig. 1 Absolute quantification of *hb* transcription reveals P1 and P2 activities in early**
906 **embryogenesis.**

907 **(a)** Schematic of the endogenous *hb* locus with two Bcd-dependent enhancers (green: distal
908 enhancer, brown: proximal enhancer) and two promoters (P1 and P2). Each promoter drives the
909 expression of a specific transcript isoform. Four smFISH probe sets were used to label different
910 regions of *hb* mRNAs. **(b)** Confocal images of two WT *Drosophila* embryos, each labeled for two
911 regions of *hb* mRNAs, Bcd protein, and DNA at nc12. Scale bars, 50 μm. Insets, magnified views

912 of anterior nuclei. Scale bars, 5 μm. Nascent signal at individual *hb* loci (in units of the number
913 of mRNA molecules) was plotted against the anterior-posterior (AP) position for different probe
914 sets. The single-locus data were binned along the AP axis (bin size: 0.05 EL, step size: 0.025
915 EL). **(c)** Percentage of active nuclei as a function of the AP position for different probe signals
916 during nc11–13. Marked region, 0.2–0.4 EL. Shadings indicate s.e.m. **(d)** Average percentage
917 of active nuclei in the position range of 0.2–0.4 EL for different probe signals during nc11–13.
918 Error bars represent s.e.m. **(e)** Nascent signal per nucleus (in units of the number of mRNA
919 molecules) as a function of the AP position for different probe sets during nc11–13. Shadings
920 indicate s.e.m. Insets, expression profiles were scaled by the maximal signal levels. **(f, g)** The
921 maximal signal level **(f)** and the boundary position **(g)** of the anterior expression domain for
922 different probe sets during nc11–13. Error bars represent s.e.m. **(h)** The Fano factor for the
923 nascent signal at individual *hb* gene loci in the position range of 0.2–0.4 EL for different probe
924 sets during nc11–13. **(c–h)** Data averaged from ⩾5 embryos for each nuclear cycle. The spatial
925 profile of each embryo was binned from the single-nucleus data (bin size: 0.05 EL, step size:
926 0.025 EL). P-values were from Student's t-test: *, $p<0.05$; **, $p<0.01$.

927

928 **Fig. 2 P1 contributes a modest yet non-negligible fraction of early *hb* transcription and**
929 **function.**

930 **(a)** Confocal image of a WT *Drosophila* embryo labeled for P1-5'UTR, P2-3'UTR, CDS of *hb*
931 mRNA, and DNA at nc13. Scale bar, 50 μm. Inset, magnified view of anterior nuclei. Scale bar,

5 μm. **(b)** Nascent P1-5'UTR, P2-3'UTR, and CDS signals per nucleus (in units of the number of mRNA molecules) were plotted against each other and fitted to a linear function. Data from a single nc12 embryo. **(c)** Parameters of the linear fit between P1-5'UTR, P2-3'UTR, and CDS signals during nc11–13. Data from one embryo in nc11, five embryos in nc12, and four embryos in nc13. See Methods and Supplementary Fig. 5a for embryo selection. Error bars represent s.e.m. **(d)** Schematic of P1 and P2 transcription. Nascent mRNAs were elongated on the gene with a constant speed and stayed on the transcription site for an extra period before being released. The P1-specific intron was spliced during or after transcription. **(e)** Average P1 and P2 contributions to the nascent CDS signal in individual nuclei in the position range of 0.2–0.4 EL during nc11–13. **(f)** The contributions of P1 and P2 to the intrinsic noise of the nascent CDS signal at individual *hb* gene loci in the position range of 0.2–0.4 EL during nc11–13. **(g, h)** Nascent mRNA signal per nucleus as a function of the AP position for *Kr* (**b**) and *kni* (**h**) in nc14. Shadings indicate s.e.m. Right: the anterior boundary positions of *Kr* and *kni* expression domains. Error bars represent s.e.m. **(e, f)** Data from the same embryos as in **(c)**. **(g, h)** Data averaged from ⩾3 nc14 embryos. The spatial profile of each embryo was binned from the single-nucleus data (bin size: 0.05 EL, step size: 0.025 EL).

**Fig. 3 Bcd activates P1 and P2 through different regulatory relations.**

**(a)** Nascent P1-5'UTR and P2-3'UTR signals per nucleus as functions of the AP position in a 1× *bcd* embryo (solid line, bin size: 0.05 EL, step size: 0.025 EL) were compared with that in the

WT (dashed line, data averaged from five embryos). Shadings indicate s.e.m. **(b)** The boundary position of the anterior expression domain for P1-5'UTR and P2-3'UTR signals in WT and 1× *bcd* embryos during nc11–13. Data averaged from ⩾5 WT embryos and ⩾3 1× *bcd* embryos for each nuclear cycle. Error bars represent s.e.m. P-values were from Student's t-test: *, $p<0.05$; **, $p<0.01$. **(c)** Nascent P1-5'UTR (upper) and P2-3'UTR (lower) signals in individual nuclei were plotted against nuclear Bcd concentration (single WT embryo, >1200 nuclei, 0.25–0.75 EL). The single-nucleus data were binned along the Bcd axis (mean ± s.e.m., bin size: 16 nM, step size: 2 nM) and fitted to Hill functions. **(d)** The Hill coefficient of the gene regulation function for different probe signals during nc11–13. Error bars represent s.e.m. P-values were from Student's t-test: *, $p<0.05$; **, $p<0.01$. **(e)** The ratios of the concentration threshold between P1-intron and CDS signals, and between P1-5'UTR and P2-3'UTR signals, during nc11–13. Error bars represent s.e.m. **(f)** The average number of Bcd molecules bound at P1- and P2-active *hb* loci in the anterior expression domain (0.2–0.35 EL) during nc11–13. Error bars represent s.e.m. Inset: an anterior nucleus labeled for P1-5'UTR and P2-3'UTR of *hb* mRNAs and Bcd protein. The enriched Bcd signal in the vicinity of the promoter (yellow and purple circles) was measured. Scale bar, 2 μm. **(g)** Bcd binding at P1- and P2-active *hb* loci as a function of nuclear Bcd concentration at nc12. Data were pooled from five embryos and binned by nuclear AP position (bin size: 0.035 EL, step size: 0.01 EL). The binned data were fitted to multi-Hill functions. Dashed lines highlight discrete binding plateaus for each promoter. Error bars represent s.e.m. **(d–g)** Data averaged from ⩾5 embryos for each nuclear cycle.

972

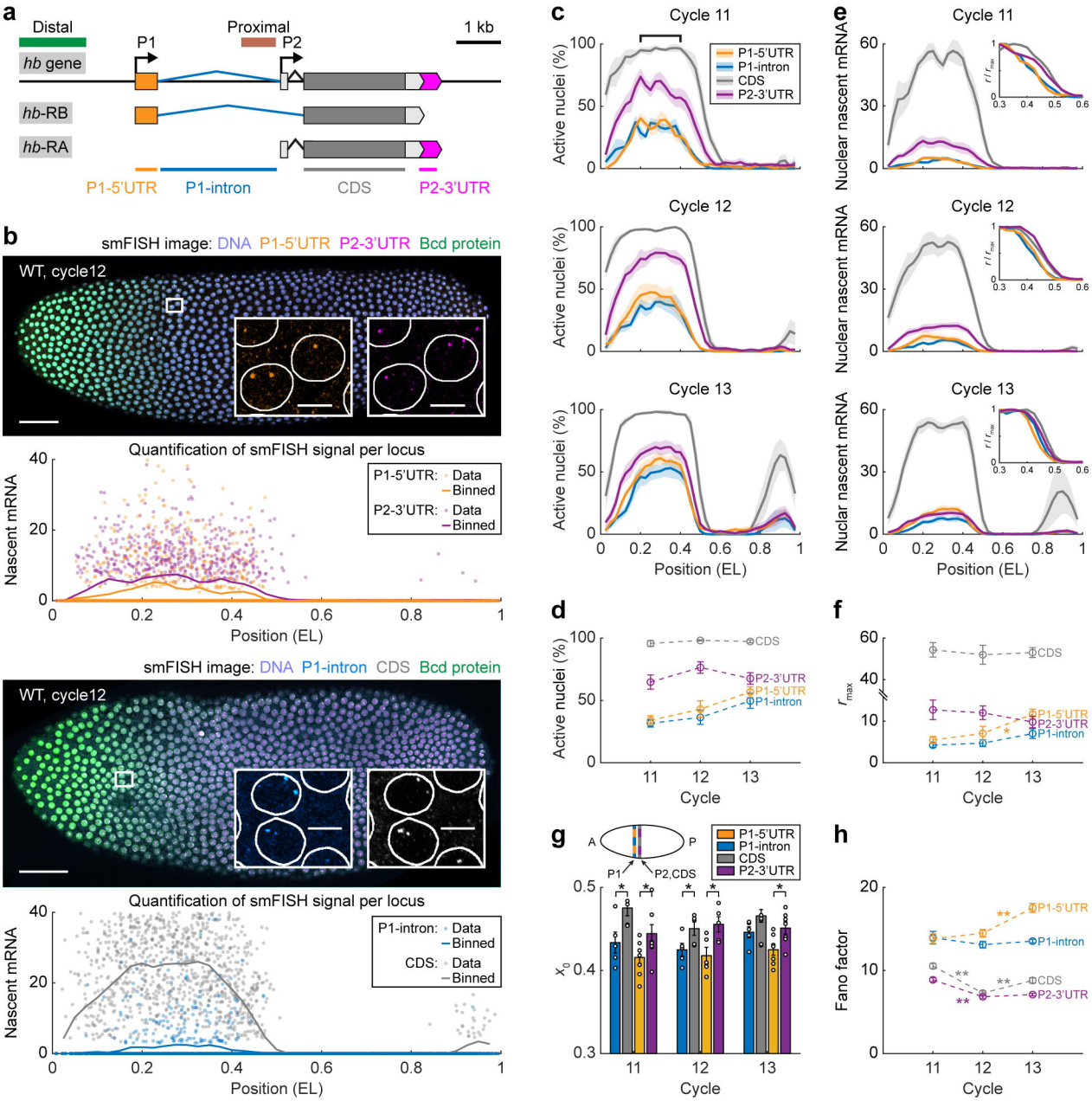
973 **Fig. 4 Two enhancers combine differently to drive P1 and P2 activation.**

974 **(a)** Schematic of *hb* reporter constructs with enhancer replacements. A construct without
975 enhancer deletion was used as a control. Two smFISH probe sets were used to label different
976 regions of reporter mRNAs: blue, P1-intron probes; red, *yellow* probes. **(b)** Confocal image of a
977 distal-enhancer-removed embryo labeled for P1-intron, *yellow*, and *hb* CDS at nc12. Scale bar,
978 50 μm. Inset, magnified view of a single anterior nucleus. Scale bar, 2 μm. **(c)** Percentage of
979 active nuclei as a function of the AP position for P1-intron and *yellow* signals in nc12 embryos of
980 different constructs. Shadings indicate s.e.m. **(d)** Average percentage of active nuclei in the
981 position range of 0.2–0.4 EL for P1-intron and *yellow* signals in different constructs during nc11–
982 13. Error bars represent s.e.m. **(e)** Nascent P1-intron and *yellow* signals per nucleus as functions
983 of the AP position in nc12 embryos of different constructs. Shadings indicate s.e.m. **(f)** The
984 maximal signal level of the anterior expression domain for P1-intron and *yellow* signals in
985 different constructs during nc11–13. Error bars represent s.e.m. **(g, h)** The boundary shift of the
986 anterior expression domain for *yellow* **(g)** and P1-intron **(h)** signals upon removing one enhancer.
987 Error bars were computed from the standard errors of boundary positions for enhancer-deleted
988 and control lines. Right: schematic of boundary shift. **(c–h)** Data compared between the distal-
989 or proximal-enhancer-removed constructs and their controls. Data averaged from ≥4 embryos
990 for each reporter construct and each nuclear cycle. The spatial profile of each embryo was

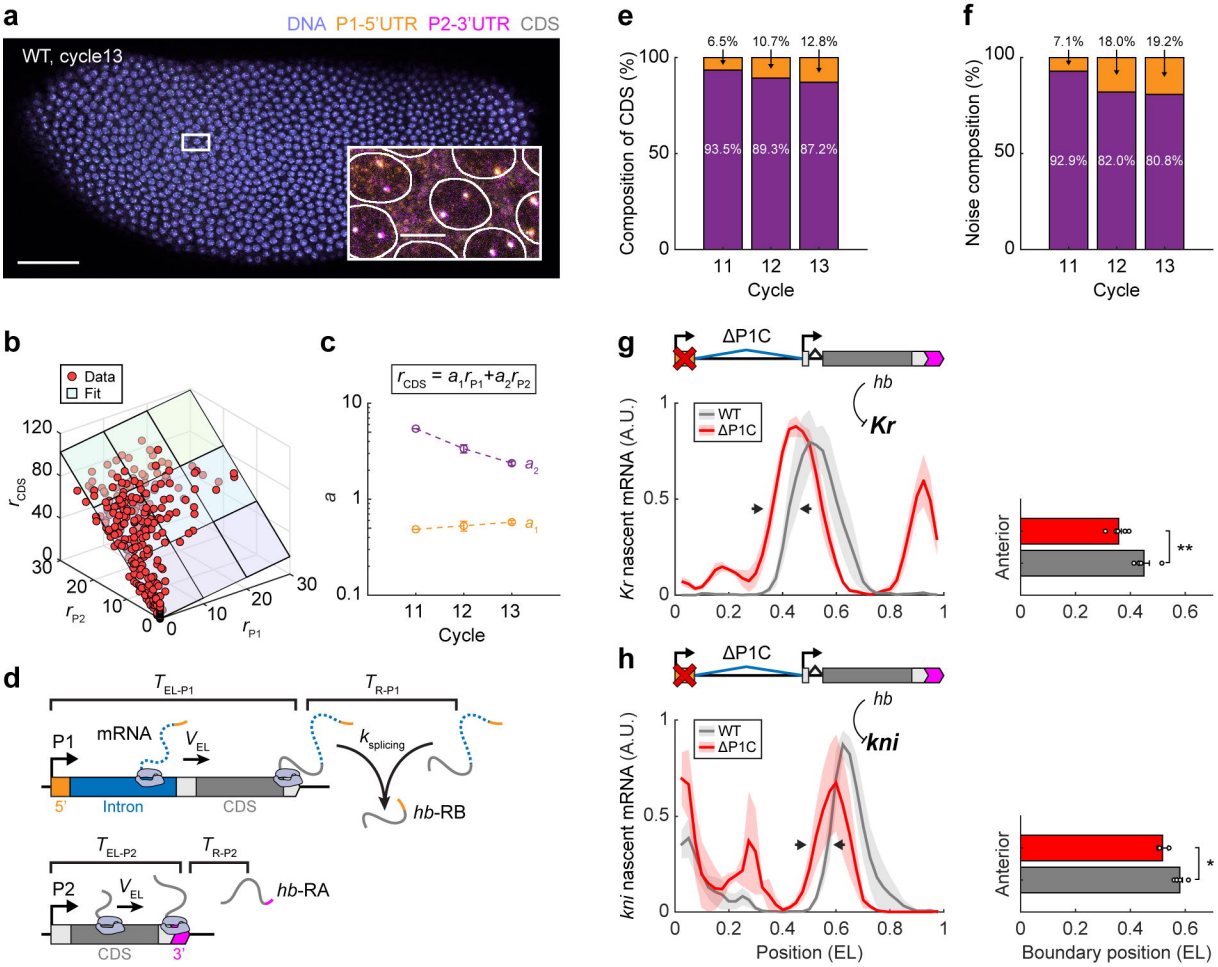
991 binned from the single-nucleus data (bin size: 0.05 EL, step size: 0.025 EL). P-values were from
992 Student's t-test: *, p<0.05; **, p<0.01.
993
994 **Fig. 5 Three-state promoter kinetics reveal a unified scheme of P1 and P2 regulation.**
995 **(a)** Histograms of nascent P1-5'UTR (upper) and P2-3'UTR (lower) signals at individual *hb* gene
996 loci in different position ranges (single embryo, the center of each position range is indicated
997 above the histogram). Each histogram was fitted to a three-state transcription model (right). **(b)**
998 Histograms of nascent P1-5'UTR (upper) and P2-3'UTR (lower) signals at active sister loci in the
999 position range of 0.2–0.4 EL. Each histogram was fitted to two Poisson distributions. Data pooled
1000 from seven embryos at nc13. Insets, images of a single anterior nucleus with active sister loci
1001 pairs. **(c)** Promoter activation rates for P1 and P2 estimated from five embryos at nc12 (binned
1002 along the AP axis, bin size: 0.1 EL, step size: 0.01 EL) were plotted against nuclear Bcd
1003 concentration and fitted to Hill functions. **(d, e)** The Hill coefficients **(d)** and concentration
1004 thresholds **(e)** of promoter activation rates for P1 and P2 during nc11–13. Error bars represent
1005 s.e.m. **(f)** The transcription initiation rates of the two active states for P1 and P2 during nc11–13.
1006 Error bars represent s.e.m. **(g)** The contributions of states 1 and 2 to P1 and P2 transcription in
1007 the position range of 0.2–0.4 EL during nc11–13. Error bars represent s.e.m. **(h)** P1 activation
1008 rates as functions of the AP position for different constructs at nc12 (≥4 embryos for each reporter
1009 construct, bin size: 0.1 EL, step size: 0.02 EL). Data were normalized and compared between
1010 the distal- or proximal-enhancer-removed constructs and the control. **(i)** Schematic of P1 and P2

1011 regulation by cooperative Bcd binding at the proximal and distal enhancers. **(j)** The joint
1012 distribution of nascent P1-5'UTR and P2-3'UTR signals at individual *hb* gene loci in the position
1013 range of 0.2–0.4 EL compared with model prediction. Data pooled from 17 embryos during nc11–
1014 13. White circles, individual loci. Color code, probability estimated from a unified model of P1
1015 and P2 transcription. **(d–g)** Data averaged from ≥5 embryos for each nuclear cycle.

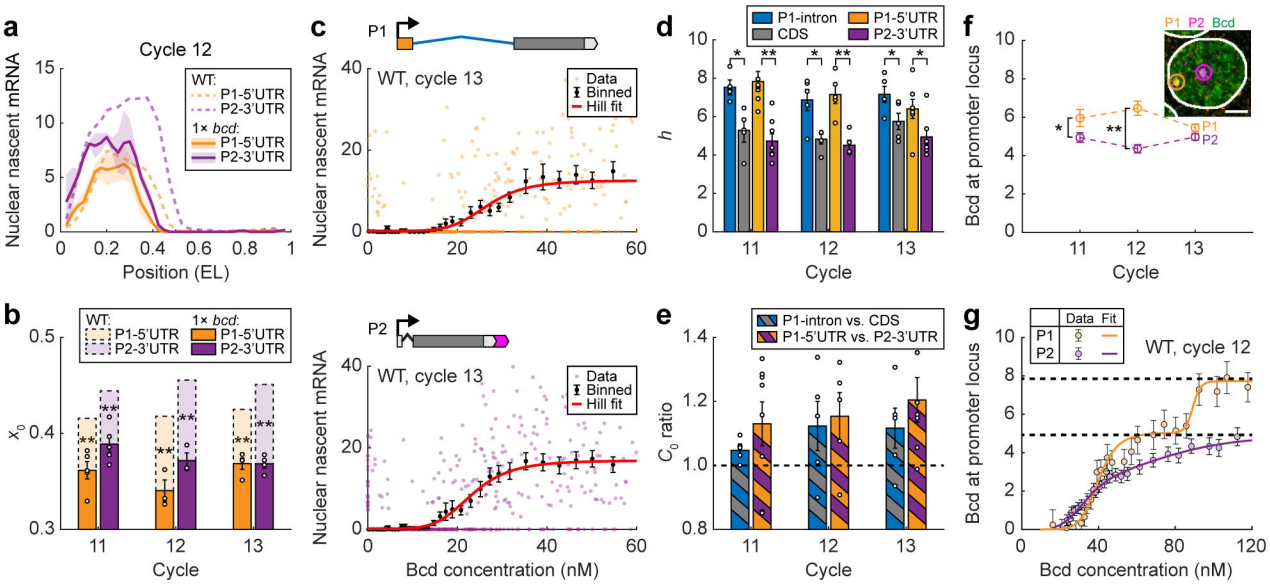
## Figure 1


a
Distal
Proximal
1 kb
P1
P2
hb gene
hb-RB
hb-RA
P1-5'UTR
P1-intron
CDS
P2-3'UTR
b
smFISH image: DNA P1-5'UTR P2-3'UTR Bcd protein
WT, cycle12
Quantification of smFISH signal per locus
Nascent mRNA
Position (EL)
P1-5'UTR: Data Binned
P2-3'UTR: Data Binned
smFISH image: DNA P1-intron CDS Bcd protein
WT, cycle12
Quantification of smFISH signal per locus
P1-intron: Data Binned
CDS: Data Binned
c
Cycle 11
Cycle 12
Cycle 13
Active nuclei (%)
Position (EL)
P1-5'UTR
P1-intron
CDS
P2-3'UTR
e
Nuclear nascent mRNA
Nuclar nascent mRNA
d
Active nuclei (%)
Cycle
f
Cycle
g
A
P
P1
P2,CDS
Cycle
h
Fano factor
Cycle

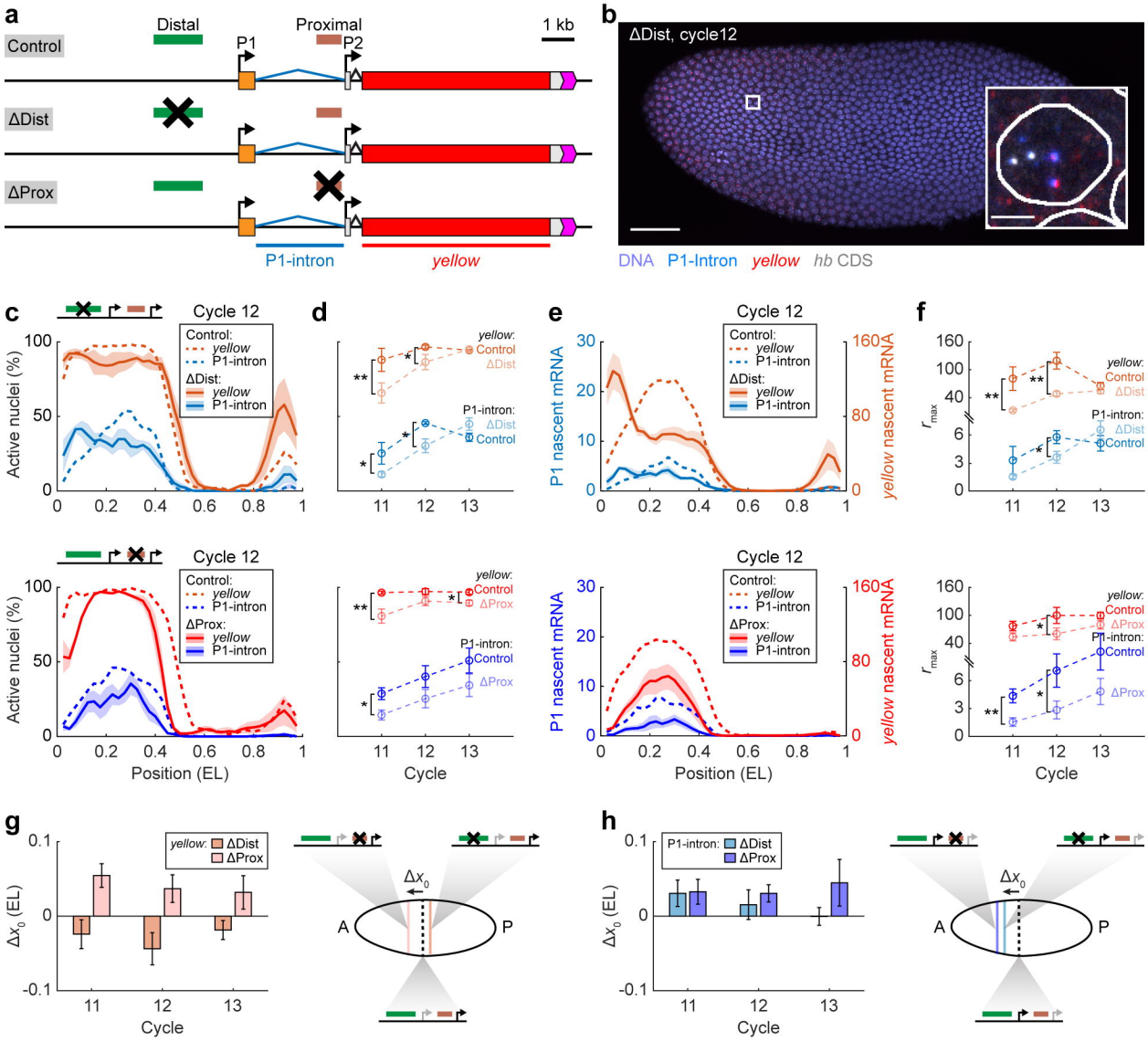
## Figure 2


a
DNA P1-5'UTR P2-3'UTR CDS
WT, cycle13
b
Data
Fit
$r_{CDS}$
$r_{P2}$
$r_{P1}$
c
$r_{CDS} = a_1 r_{P1} + a_2 r_{P2}$
a
$a_2$
$a_1$
Cycle
d
$T_{EL-P1}$
$T_{R-P1}$
P1
mRNA
$V_{EL}$
$k_{splicing}$
5'
Intron
CDS
hb-RB
$T_{EL-P2}$
$T_{R-P2}$
P2
hb-RA
3'
e
Composition of CDS (%)
6.5%
10.7%
12.8%
93.5%
89.3%
87.2%
Cycle
f
Noise composition (%)
7.1%
18.0%
19.2%
92.9%
82.0%
80.8%
Cycle
g
ΔP1C
hb
Kr
WT
ΔP1C
Kr nascent mRNA (A.U.)
Anterior
**
h
ΔP1C
hb
kni
WT
ΔP1C
kni nascent mRNA (A.U.)
Position (EL)
Anterior
*
Boundary position (EL)

## Figure 3

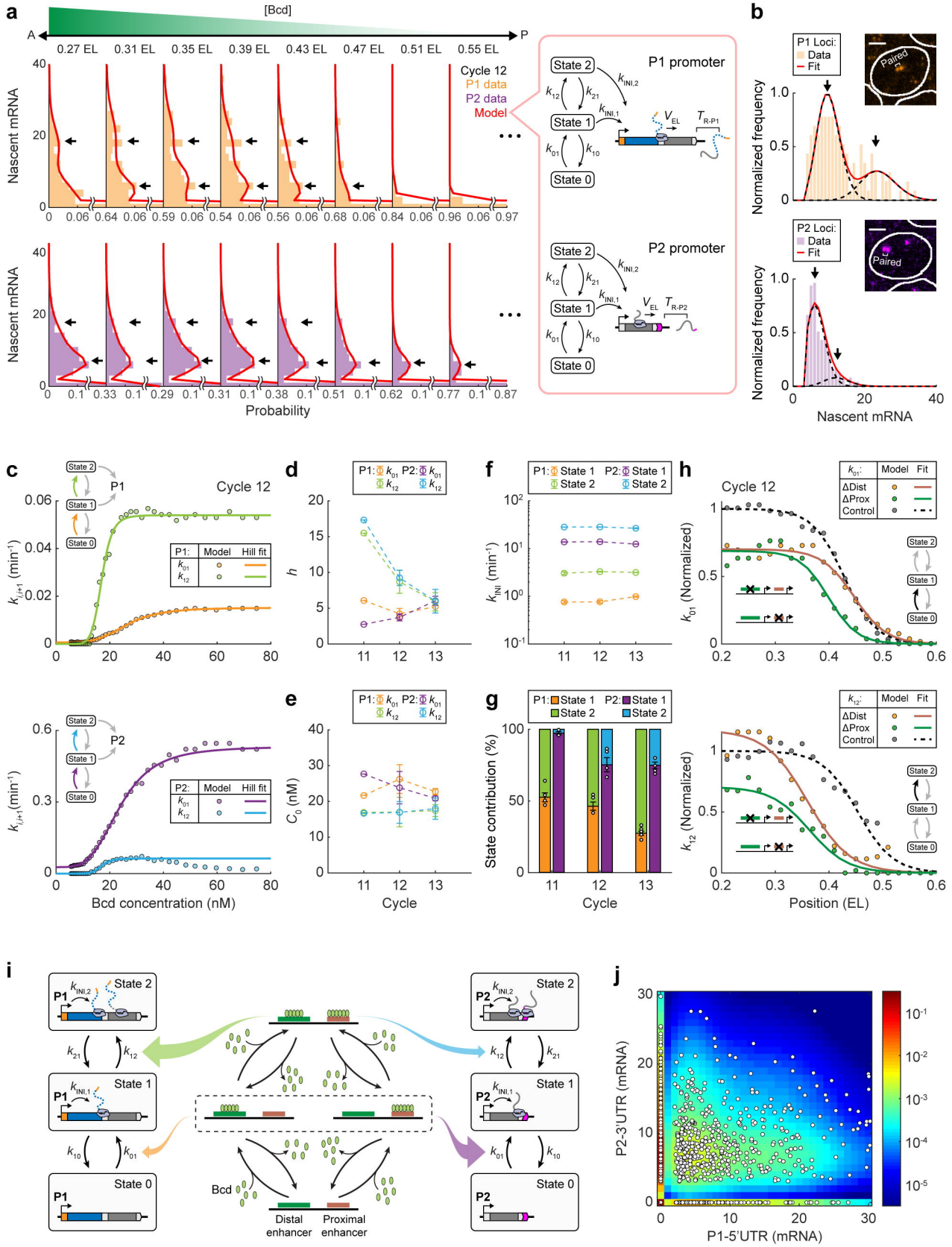

a
Cycle 12
Nuclear nascent mRNA
Position (EL)
WT:
P1-5'UTR
P2-3'UTR
1× bcd:
P1-5'UTR
P2-3'UTR
b
$x_0$
Cycle
c
P1
WT, cycle 13
Data
Binned
Hill fit
Nuclear nascent mRNA
P2
Bcd concentration (nM)
d
P1-intron
CDS
P1-5'UTR
P2-3'UTR
h
Cycle
e
P1-intron vs. CDS
P1-5'UTR vs. P2-3'UTR
$C_0$ ratio
Cycle
f
P1 P2 Bcd
Bcd at promoter locus
Cycle
g
Data Fit
P1
P2
WT, cycle 12
Bcd at promoter locus
Bcd concentration (nM)

## Figure 4



a
Distal
Proximal
1 kb
Control
P1
P2
ΔDist
ΔProx
P1-intron
yellow
b
ΔDist, cycle12
DNA
P1-Intron
yellow
hb CDS
c
Cycle 12
Control:
yellow
P1-intron
ΔDist:
yellow
P1-intron
Active nuclei (%)
Position (EL)
d
yellow:
Control
ΔDist
P1-intron:
ΔDist
Control
Cycle
e
P1 nascent mRNA
yellow nascent mRNA
f
$r_{max}$
g
yellow: ΔDist
ΔProx
$\Delta x_0$ (EL)
A
P
h
P1-intron: ΔDist
ΔProx

## Figure 5


a
[Bcd]
A
P
0.27 EL
0.31 EL
0.35 EL
0.39 EL
0.43 EL
0.47 EL
0.51 EL
0.55 EL
Nascent mRNA
Cycle 12
P1 data
P2 data
Model
Probability
P1 promoter
P2 promoter
State 2
State 1
State 0
b
P1 Loci:
Data
Fit
Paired
Normalized frequency
Nascent mRNA
P2 Loci:
c
Cycle 12
P1
P2
Bcd concentration (nM)
d
h
e
$C_0$ (nM)
Cycle
f
$k_{INI}$ (min$^{-1}$)
g
State contribution (%)
h
$k_{01}$ (Normalized)
$k_{12}$ (Normalized)
Position (EL)
ΔDist
ΔProx
Control
i
Bcd
Distal enhancer
Proximal enhancer
j
P2-3'UTR (mRNA)
P1-5'UTR (mRNA)