

# 1 **Concentration Dependent Chromatin States Induced by the Bicoid Morphogen**

## 2 **Gradient**

3 Colleen E. Hannon<sup>1</sup>, Shelby A. Blythe<sup>1</sup> and Eric F. Wieschaus<sup>1\*</sup>

4 <sup>1</sup>Department of Molecular Biology/Howard Hughes Medical Institute, Princeton University, Princeton, NJ  
5 08544, USA

6 \*Correspondence: [efw@princeton.edu](mailto:efw@princeton.edu)

## 7 **ABSTRACT**

8         In *Drosophila*, graded expression of the maternal transcription factor Bicoid (Bcd)  
9 provides positional information to activate target genes at different positions along the  
10 anterior-posterior axis. We have measured the genome-wide binding profile of Bcd  
11 using ChIP-seq in embryos expressing single, uniform levels of Bcd protein, and  
12 grouped Bcd-bound targets into four classes based on occupancy at different  
13 concentrations. By measuring the biochemical affinity of target enhancers in these  
14 classes *in vitro* and genome-wide chromatin accessibility by ATAC-seq, we found that  
15 the occupancy of target sequences by Bcd is not primarily determined by Bcd binding  
16 sites, but by chromatin context. Bcd drives an open chromatin state at a subset its  
17 targets. Our data support a model where Bcd influences chromatin structure to gain  
18 access to concentration-sensitive targets at high concentrations, while concentration-  
19 insensitive targets are found in more accessible chromatin and are bound at low  
20 concentrations.

## 21 **INTRODUCTION**

22         During embryonic development, multicellular organisms must generate the  
23 patterned tissues of an adult organism from a single undifferentiated cell. This process

24 requires highly regulated control of gene expression, both in developmental time and at  
25 reproducible positions in an embryo. These complex gene regulatory networks are  
26 controlled by systems of transcription factors, which bind to DNA and control the  
27 expression of genes required for development (Levine and Davidson, 2005). In early  
28 *Drosophila melanogaster* embryos, the Bicoid transcription factor forms an anterior-to-  
29 posterior protein gradient the embryo (Driever and Nüsslein-Volhard, 1988b). Bcd  
30 functions as transcriptional activator to pattern the embryo, binding to target gene  
31 enhancers and activating gene expression at distinct positions along the AP axis,  
32 corresponding to different concentrations of Bcd protein (Driever and Nüsslein-Volhard,  
33 1988a; Struhl et al., 1989).

34         Recent studies of Bcd function suggest that its interaction with its targets may be  
35 more complex than the simple concentration-dependent activation originally proposed  
36 for morphogen gradients (Wolpert, 1969). In the absence of a strong Bcd gradient,  
37 embryos still exhibit patterned expression of Bcd target genes, and these genes can be  
38 activated at lower concentrations of Bcd than these nuclei would be exposed to in a  
39 wild-type embryo (Chen et al., 2012; Liu et al., 2013; Ochoa-Espinosa et al., 2009).  
40 While changing Bcd dosage shifts cell fates, the shifts deviate quantitatively from those  
41 expected of strict concentration dependence, especially as expression patterns are  
42 assayed progressively later during development (Liu et al., 2013). These studies have  
43 consequently raised doubts about the extent to which the local concentration of Bcd  
44 along its gradient determines the spatial patterns of target gene expression in the  
45 embryo.

46           While these studies argue against a strict application of the morphogen  
47 hypothesis for Bcd, the patterned expression of target genes is also influenced by other  
48 maternal patterning systems and interactions among the Bcd targets themselves (Chen  
49 et al., 2012; Jaeger, 2010; Löhr et al., 2009). Chen, *et al.* have shown that the posterior  
50 boundaries of Bcd target genes are positioned by a system of repressors including  
51 Runt, Krüppel, and Capicua. This work suggests that the Bcd gradient does not directly  
52 establish expression domains of its target genes but rather is just one player in a  
53 network of patterning genes that influence cell fates in the embryo. However, using  
54 target gene expression as a metric for Bcd function does not address how information  
55 from the Bcd gradient initially establishes distinct cell fates.

56           Part of the difficulty in evaluating direct roles of the Bcd gradient arises from the  
57 unknown nature of the molecular mechanism by which Bcd establishes concentration  
58 thresholds different positions along the gradient. A simple model of the positioning of  
59 Bcd target genes predicts that *cis*-regulatory elements of different genes respond to  
60 different concentrations of Bcd. Genes in the anterior would have low affinity Bcd  
61 binding sites and could therefore only be activated by high Bcd concentrations, whereas  
62 genes expressed in more posterior positions would have higher affinity binding sites  
63 (Driever et al., 1989b). Direct measurements of Bcd binding affinity have been  
64 conducted *in vitro* using DNA probes (Burz et al., 1998; Gao and Finkelstein, 1998; Ma,  
65 1996; Ma et al., 1996) and have demonstrated that Bcd is able to bind cooperatively to  
66 achieve sharp concentration thresholds. While these measurements lend some support  
67 to a simple affinity model, little correlation has been shown between predicted binding  
68 site affinity and AP position of gene expression (Ochoa-Espinosa et al., 2005; Segal et

69 al., 2008). However, neither *in vitro* measurements of Bcd binding nor computational  
70 predictions of binding sites can capture interactions between Bcd and its target  
71 enhancers in the context of local chromatin structure.

72 Using high throughput sequencing approaches, we measured *in vivo* genome-  
73 wide Bcd-DNA binding and chromatin accessibility in transgenic embryos expressing  
74 different concentrations of uniform Bcd protein. These data reveal distinct classes of  
75 enhancers that differ in their sensitivity to Bcd concentration. We find that these classes  
76 differ both in the DNA binding motifs that they contain and in their local chromatin  
77 accessibility. We also find that Bcd influences the accessibility of a subset of its target  
78 enhancers, primarily at highly concentration-sensitive enhancers that drive gene  
79 expression in the anterior of the embryo. This leads us to a model in which target  
80 enhancers throughout the genome have a broad range of sensitivities for Bcd protein,  
81 and can therefore respond to a range of Bcd concentrations along the gradient.  
82 However, rather than arising from differences in Bcd binding site composition, these *in*  
83 *vivo* interactions are chromatin context-dependent, and Bcd influences the chromatin  
84 structure of its target enhancers.

## 85 **RESULTS**

### 86 **Bicoid target gene expression boundaries are influenced by other maternal** 87 **factors, but its physical interaction with enhancers is not**

88 To investigate the mechanism whereby Bcd functions to pattern the AP axis, we  
89 performed chromatin immunoprecipitation followed by high throughput sequencing  
90 (ChIP-seq) to determine the genome-wide binding profile of Bcd to its targets. We  
91 performed the ChIP-seq experiments on embryos expressing GFP-tagged Bcd in a *bcd*

92 null mutant background that were staged precisely in nuclear cycle 14 (NC14), and  
93 established a list of robust and reproducible list of 1,027 peak Bcd binding regions (see  
94 Supplemental File 1 and Experimental Procedures). These peaks successfully identify  
95 63 out of 66 of the previously identified Bcd target enhancers (Chen et al., 2012) and  
96 overall associate with enhancers whose expression patterns span broadly across the  
97 AP axis.

98 As a transcriptional regulator, Bcd activates the expression of a subset of its  
99 targets whose expression domains are predominantly located in the anterior half of the  
100 embryo. In *bcd* mutant embryos, such targets are not expressed. For example, the gap  
101 genes *buttonhead* (*btd*) and *knirps* (*kni*) have anterior expression domains that are not  
102 present in *bcd* mutant embryos (Figure 1A). The posterior *kni* expression domain,  
103 however, is expressed in *bcd* embryos, albeit with shifted positional boundaries. These  
104 distinct domains of *kni* expression are controlled by separate enhancer elements  
105 (Pankratz et al., 1992; Schroeder et al., 2004), both of which are bound by Bcd *in vivo*  
106 (Supplemental File 1). In the absence of all maternal AP patterning inputs (*bicoid nanos*  
107 *hunchback torsolike* quadruple mutants), the *kni* posterior enhancer is not expressed.  
108 However, in embryos where Bcd is the sole source of maternal patterning information  
109 (*nanos hunchback torsolike* triple mutants), the *kni* posterior domain is expressed with a  
110 near wild-type anterior expression boundary (Figure 1A and Figure 1 Figure  
111 Supplement 1A). The *kni* posterior domain therefore represents a second class of Bcd  
112 target gene, which depends on Bcd to determine the position of its expression but does  
113 not demonstrate an absolute requirement of Bcd for transcriptional activation.

114 Both classes of Bcd target genes receive positional cues both from Bcd and from  
115 other patterning systems. We considered the possibility that, given their influence on the  
116 expression domains of Bcd target genes, the posterior and terminal patterning systems  
117 may impact Bcd binding to its target enhancers in different nuclei along the AP axis. We  
118 therefore tested whether loss of the posterior and terminal systems (*nanos* and  
119 *torsolike*) would alter the Bcd ChIP-seq profile. We used the statistical package EdgeR  
120 (Robinson et al., 2010) to test for differential Bcd binding between wild-type and *nanos*  
121 *torsolike* embryos and found that we could not detect any significant change in binding  
122 at any of these 1,027 regions (Figure 1 Figure Supplement 1B). Therefore, although the  
123 expression domains of Bcd target genes are ultimately influenced by inputs from other  
124 AP patterning systems, the physical interaction of Bcd with the DNA in the enhancers of  
125 these genes occurs independently of other maternal AP patterning inputs.

### 126 **Embryos expressing Bcd uniformly show developmental fates reflecting the** 127 **concentration of Bcd**

128 We set out test whether incremental changes in Bcd concentration along the  
129 gradient can be read out directly at the level of binding to target enhancers. Due to the  
130 graded distribution of Bcd, each nucleus along the AP axis is exposed to a different  
131 concentration of the protein. To measure the Bcd binding state at individual  
132 concentrations, we performed ChIP-seq on embryos expressing Bcd at single, uniform  
133 concentrations in every nucleus along the AP axis. Several previous studies have  
134 included genetic manipulations in which the Bcd gradient has been flattened to assess  
135 its activity independently of its distribution (Chen et al., 2012; Driever and Nüsslein-  
136 Volhard, 1988a; Löhr et al., 2009; Ochoa-Espinosa et al., 2009). However, genetically

137 disrupting the gradient does not result in a total flattening, and transgenic approaches to  
138 date have not allowed for precise and reproducible control over the level of expression  
139 of the flattened Bcd. We therefore generated transgenic lines expressing GFP tagged  
140 Bcd in which the endogenous 3'UTR responsible for graded localization is flanked by  
141 FRT sites that allow it to be replaced with the unlocalized *spaghetti squash* 3'UTR. To  
142 generate different expression levels of uniform Bcd, we coupled transgenes to different  
143 maternally active promoters that yield embryos in which individual uniform Bcd  
144 concentration approximates single positions along the gradient (see Figure 1 Figure  
145 Supplement 1C).

146 To determine the expression levels of the uniform lines, we imaged GFP  
147 fluorescence in live embryos expressing either uniform or graded GFP-Bcd (Gregor et  
148 al., 2007a) (Figure 1B). The endogenous *bcd* promoter drives a level of uniformly  
149 expressed Bcd equivalent to that measured at approximately 65% egg length of the  
150 wild-type gradient. The *matrimony* (*mtrm*) and  $\alpha$ *Tubulin67C* ( $\alpha$ *Tub67C*) promoters drive  
151 expression levels corresponding to approximately 45% and 25% egg length,  
152 respectively. For simplicity, we refer to the uniform lines as low (*bcd* promoter), medium  
153 (*mtrm* promoter), and high ( $\alpha$ *Tub67C* promoter). (See also Figure 1 Figure Supplement  
154 1D)

155 Uniform expression of Bcd confers gene expression profiles and developmental  
156 programs representative of distinct positions along the AP axis. The head gap gene  
157 *buttonhead* (*btd*) is expressed in an anterior stripe in wild-type embryos (Figure 1A), but  
158 expands to fill the entire middle of the embryo at the highest level of uniform Bcd (Figure  
159 1C). At the medium level, the Btd anterior stripe is duplicated at the posterior, and at the

160 lowest Bcd level, it is not expressed. The gap gene *knirps*, which is expressed in an  
161 anterior domain and a posterior stripe, shows a duplication of its anterior domain in the  
162 posterior in high uniform Bcd embryos. There is also a weaker duplication at the  
163 medium Bcd level. There is no apparent anterior expression at the lowest level, but an  
164 expanded posterior stripe is present. The gene expression patterning we observe in the  
165 presence of uniform Bcd likely result from the activity of additional maternal patterning  
166 cues (*nanos* and *torso*) as well as interactions between the Bcd target genes  
167 themselves. The concentration-dependent activity of uniform Bcd is also apparent in  
168 cuticle preparations of embryos expressing the transgenic constructs. The transgenic  
169 constructs specify increasingly anterior structures along larval body plan as the  
170 concentration of Bcd increases (Figure 1D). These effects on the body plan indicate that  
171 the uniform Bcd transgenes are capable of specifying cell fates that reflect their relative  
172 expression levels.

### 173 **Bcd binding to genomic targets is concentration dependent**

174 We next determined genome-wide Bcd binding profiles at each individual  
175 concentration by ChIP-seq and used these measurements to assign each of the 1,027  
176 peak regions to classes distinguished by their degree of concentration-dependent Bcd  
177 binding (Figure 2A). Using EdgeR (Robinson et al., 2010), we selected peak regions  
178 that exhibited statistically significant ( $FDR \leq 0.05$ ) differences in binding by performing  
179 pairwise exact tests between the three uniform Bcd concentrations. This yielded four  
180 different classes of peaks, one concentration-insensitive class, and three classes with  
181 increasing sensitivity to Bcd concentration.



182           The Concentration-Insensitive peak class (n = 143) shows no significant  
183 differences between any of the concentrations of uniform Bcd we tested. Concentration-  
184 Sensitive III peaks (n = 593) are significantly reduced in binding between the highest  
185 and lowest Bcd concentrations, but reductions are not significant between high and  
186 medium, or medium and low. Concentration-Sensitive II peaks (n = 138) are significantly  
187 reduced in binding at the lowest Bcd level compared to either the medium or the high  
188 levels. Finally, Concentration-Sensitive I peaks (n = 152) are significantly reduced in  
189 binding at both the medium and the low Bcd levels compared to the highest level  
190 (Figure 2A). These different groups suggest that Bcd binds differentially to target  
191 enhancers at specific concentrations, and furthermore that certain subsets of enhancers  
192 are bound only in anterior nuclei whereas others are bound broadly across the entire AP  
193 axis.

194           Although 63 out of 66 previously characterized Bcd-dependent enhancers are  
195 identified in our ChIP peaks, the majority of the 1,027 peaks identified have not been  
196 extensively examined. Within the set of known Bcd targets, there is strong correlation  
197 between position of expression and the associated Bcd sensitivity class (Figure 2A). To  
198 extend this observation to previously uncharacterized Bcd target enhancers, we queried  
199 the Fly Enhancer resource generated from the Vienna Tile GAL4 reporter library (Kvon  
200 et al., 2014). The Fly Enhancer collection is a library of candidate enhancer DNA  
201 fragments driving expression of GAL4 that covers 13.5% of the non-coding genome.  
202 Each fragment's expression pattern has been measured and scored by developmental  
203 stage. A total of 293 enhancer candidates overlap with at least one peak in our data set.  
204 Of these, 163 drive gene expression in stage 4-6 (which includes NC14), and these

205 active enhancers overlap with a total of 151 (14.7%) of the Bcd-peaks. The remaining  
206 overlapping fragments either are active later in development (75), or are not functional  
207 (55). Given the large fraction of the queried enhancers that are active during early  
208 development (163 out of 293 overlapping enhancers), it remains possible that a similar  
209 fraction of the of the 876 peaks (77.2%) that do not overlap with the Fly Enhancer  
210 candidates may correspond to enhancers active in the early embryo.

211 The Bcd sensitivity classes are predictive of the expression domains of  
212 associated enhancer fragments. Enhancers overlapping with both the Concentration-  
213 Sensitive I and II classes drive expression in anterior regions of the embryo, with the  
214 Concentration-Sensitive III and Concentration-Insensitive classes driving broad and  
215 posterior expression, respectively (Figure 2B). This indicates that our classifications of  
216 the Bcd-bound peaks reflect unique groups of Bcd targets with differing abilities to bind  
217 Bcd protein and consequently activate gene expression in different positions along the  
218 AP axis. The boundary positions of anteriorly expressed Bcd targets may be refined at  
219 the transcriptional level by interactions with opposing gradients of repressors like Runt  
220 (Chen et al., 2012), that are Bicoid targets themselves. We addressed whether such  
221 repression could account for the restricted expression of Concentration-Sensitive I and  
222 II targets to anterior regions of the embryo by examining whether they were enriched for  
223 binding of such repressors. By comparison with genome-wide binding profiles of  
224 transcription factors in the BDTNP ChIP database (Li et al., 2008; MacArthur et al.,  
225 2009), we instead find that the factors associated with each Bcd peak class are  
226 generally those whose expression patterns overlap with the average expression  
227 domains of each class (Figure 2 Figure Supplement 1A). Peaks in the Concentration-

228 Sensitive I class, for example, are enriched for binding of the terminal gap gene  
229 Hucklebein whereas those in the Concentration-Sensitive II class are enriched for  
230 Krüppel and Giant binding. However, we find no evidence that the Bcd sensitivity  
231 classes are predominantly defined by repressive interactions.

232 **Sequence composition of ChIP sensitivity classes does not account for *in vivo***  
233 **sensitivity to Bcd concentration**

234 We next wanted to determine whether the Bcd-bound regions in each sensitivity  
235 class differ at the level of DNA sequence. *In vitro*, Bcd binds with high affinity to the  
236 consensus 5'-TCTAATCCC-3', and that variations on this consensus sequence  
237 constitute weak binding sites (Burz et al., 1998; Driever and Nüsslein-Volhard, 1989;  
238 Driever et al., 1989b). If the affinity of a given enhancer for Bcd were encoded primarily  
239 at the level of its DNA sequence, we would expect to see a higher representation of  
240 strong Bcd binding sites in the less sensitive classes, and weaker sites in the more  
241 sensitive classes. To test this, we performed *de novo* motif discovery using the RSAT  
242 peak-motifs algorithm (Thomas-Chollier et al., 2012; 2008). We identified the top motifs  
243 in the entire Bcd ChIP peak list, ranked by their e-value, and found that the top three  
244 most highly ranked motifs were the consensus binding site for the proposed pioneer  
245 factor Zelda (Zld) (Bosch et al., 2006; De Renzis et al., 2007; Harrison et al., 2011; Nien  
246 et al., 2011), and a strong (TAATCC) and weak (TAAGCC) Bcd binding site (Figure 2C).  
247 We next calculated the frequency with which these motifs appear in each peak, and  
248 tested for enrichment between sensitivity classes by permutation test (Figure 2D). We  
249 found that despite their failure to bind Bcd at low concentrations, the Concentration-  
250 Sensitive I and II classes are enriched for both the strong and weak Bcd sites relative to

251 the peak set as a whole. Given our result that these classes drive expression primarily  
252 in the anterior of the embryo (Figure 2B), the higher density of Bcd binding sites in these  
253 enhancers contrasts with previous studies that have found little correlation between  
254 number of binding sites and position of gene expression (Ochoa-Espinosa et al., 2005).  
255 This difference likely reflects the larger sample size used in our study, as well as our  
256 method for classifying Bcd bound peaks. The Concentration-Sensitive III class did not  
257 contain an enrichment of any site over the total peak set. The Concentration-Insensitive  
258 class, however, showed a higher prevalence of the Zld binding site relative to the total  
259 peak set than any other class.

260         These results indicate that, in contrast to a binding site affinity model for Bcd  
261 function, Bcd target enhancers that behave as concentration-sensitive and -insensitive  
262 *in vivo* are not distinguished by their representation of strong versus weak Bcd binding  
263 sites, confirming previous studies (Ochoa-Espinosa et al., 2005). In further support of  
264 this concept, we found little correlation between *in vitro* binding affinity by  
265 electrophoretic mobility shift assay and the *in vivo* binding properties we observe by  
266 ChIP for a selected subset of peaks (Figure 2 Figure Supplement 2). At the level of  
267 sequence composition, they instead appear to differ in their balance of Bcd and Zld  
268 binding sites. Although both strong and weak Bcd sites and Zld sites are enriched in the  
269 Bcd ChIP peaks as a whole, there is a bias toward both Bcd sites in peaks that show  
270 concentration-sensitive binding properties by ChIP-seq and a bias toward Zld sites in  
271 the concentration-insensitive peaks. Zld, a ubiquitously expressed early embryonic  
272 transcription factor, has been implicated in chromatin remodeling prior to zygotic  
273 genome activation (Harrison et al., 2011; Nien et al., 2011; Sun et al., 2015). The

274 predominance of Zld motifs over Bcd motifs in the Concentration-Insensitive class  
275 suggests that *in vivo* chromatin structure also plays a role in the sensitivity of a given  
276 target to transcription factor concentration in the context of the developing embryo.  
277 Taken together, these findings suggest that the chromatin context of an enhancer may  
278 play a greater role in its overall affinity for a transcription factor *in vivo* than the  
279 sequences of the binding sites that it contains. We therefore set out to test the  
280 hypothesis that sensitivity classes are distinguished at the level of chromatin structure.  
281 **Bcd is required for chromatin accessibility at a subset of concentration-sensitive**  
282 **target sites**

283       To measure genome-wide patterns of chromatin accessibility and nucleosome  
284 positioning, we performed ATAC-seq (Buenrostro et al., 2015) on single wild-type  
285 embryos precisely staged at 12 minutes after the onset of NC14, and identified 13,266  
286 peaks of chromatin accessibility (see Experimental Procedures). Of the 1,027 Bcd-  
287 bound regions identified by ChIP-seq, 855 (83.3%) of them overlap with ATAC-seq  
288 peaks.

289       Given Zelda's role in influencing chromatin accessibility and the presence of its  
290 binding sites at Bcd-bound regions of genome, we measured the effect of Zld on  
291 accessibility at Bcd sites by ATAC seq (Figure 3A). Of the total 13,226 accessible  
292 regions at NC14, 2,675 (20.2%) show a significant reduction in accessibility in *zld*  
293 mutant embryos. This fraction is higher in Bcd-bound peaks; 402 (39.1%; or 379  
294 [44.3%] of 855 the Bcd peaks that overlap with ATAC open regions, see Table S5)  
295 show reduced accessibility in *zld* mutants, indicating that Bcd bound regions are more  
296 likely to be dependent on Zld for their accessibility than the genome as a whole.

297 However, the Zld-dependent peaks are distributed across each sensitivity class  
298 determined by ChIP, with no particular class being significantly more Zld-dependent.  
299 This contrasts with the distribution of binding sites in the peak classes, which revealed  
300 that the Concentration-Insensitive peaks were more likely to contain Zld binding sites.  
301 These results suggest that while Zld contributes to the accessibility of a subset of Bcd  
302 target gene enhancers, it is unlikely to determine the differential concentration sensitivity  
303 of Bcd peaks as a whole.

304         Given the enrichment for both strong and weak Bcd binding sites in the  
305 Concentration-Sensitive I and II classes, we next examined the impact of Bcd protein  
306 itself on chromatin accessibility by ATAC seq (Figure 3A). In *bcd* mutants, 326 (2.4%) of  
307 the 13,266 open regions in wild-type embryos show significantly reduced accessibility  
308 accompanied by increased nucleosome occupancy in those same regions (Figure 3A  
309 and B). These regions are therefore either directly or indirectly dependent on Bcd for  
310 their accessibility. More strikingly, 132 (12.9%) of the 1,027 Bcd ChIP-seq peaks show  
311 reduced accessibility in the absence of Bcd and likely represent regions where Bcd's  
312 impact is direct. These regions dependent on Bcd for accessibility are significantly  
313 enriched for peaks in the Concentration-Sensitive I and II classes (32.9% and 31.9% of  
314 each class, with Fisher's exact test P-values of  $4.29 \times 10^{-12}$  and  $1.37 \times 10^{-10}$ , respectively).  
315 In contrast, the Concentration-Sensitive II and Concentration-Insensitive classes are  
316 both significantly underrepresented (6.07% and 0.7% and P-values =  $7.88 \times 10^{-13}$  and  
317  $2.29 \times 10^{-8}$ ) (Figure 3B). This suggests that Bcd binding influences chromatin  
318 accessibility preferentially at a subset of highly concentration-sensitive enhancers.

319           Because the Concentration-Sensitive I and II classes are bound primarily at high  
320 Bcd concentrations, Bcd's effects on chromatin accessibility at these targets likely  
321 occurs only in anterior regions of the embryo. In support of this, we find that chromatin  
322 accessibility at Bcd-dependent, concentration-sensitive targets is responsive to Bcd  
323 concentration. Expressing uniform Bcd confers accessibility to peaks that are not  
324 accessible in *bcd* mutant embryos (Figure 3B). The degree of chromatin accessibility  
325 conferred by Bcd correlates positively with the concentration of uniform Bcd expressed  
326 (Figure 3B). This observation, along with the overrepresentation of the Concentration-  
327 Sensitive I and II classes in the Bcd-dependent peaks, suggests that Bcd influences the  
328 chromatin state of these targets primarily at the high concentrations found in the anterior  
329 of the embryo.

330           A second feature that distinguishes the chromatin structure of Bcd binding sites  
331 is the presence of DNA sequences favorable for nucleosome occupancy (Segal et al.,  
332 2006). Bcd bound regions in wild type embryos are generally depleted of nucleosomes  
333 (Figure 4A). However, predicting nucleosome positioning sequences using the NuPoP  
334 algorithm (Xi et al., 2010) suggests that the Concentration-Sensitive I and II Bcd  
335 enhancer classes are more likely to bind nucleosomes than the Concentration-Sensitive  
336 III and Concentration-Insensitive classes. (Figure 4B). The contrast between predicted  
337 occupancy and observed depletion suggests that these regions are actively restructured  
338 for Bcd and other transcription factors to bind. The increased nucleosome preference of  
339 the more Concentration-Sensitive peaks, combined with the observation that these sites  
340 become occupied by nucleosomes in *bcd* mutants suggests a model where Bcd, either  
341 directly or in combination with cofactors is able to direct chromatin remodeling events,

342 which may play a significant role in distinguishing concentration-sensitive and -  
343 insensitive targets. Additionally, we find that Bcd-bound regions that are dependent on  
344 either Zld or Bcd for their accessibility are more likely to have a higher nucleosome  
345 preference than regions that are independent of both factors (Figure 4C). This further  
346 suggests that Bcd is able to overcome a high nucleosome barrier in a manner similar to  
347 Zld (Sun et al., 2015) at a subset of its target enhancers.

348         These effects of chromatin accessibility impact the availability of sequence motifs  
349 for binding Bcd. In wild type embryos, there is a gradual increase in average motif  
350 accessibility from high to low sensitivity, and this difference becomes more pronounced  
351 in *bcd* mutant embryos (Figure 4D), consistent with a role for Bcd in driving changes in  
352 accessibility at more sensitive sites in a concentration dependent manner. This is also  
353 evident at the level of nucleosome organization. Calculating the fraction of motifs that  
354 overlap nucleosomes in either wild type or *bcd* mutant chromatin conformations, we find  
355 that whereas on average across all sensitivity classes  $55 \pm 2\%$  of Bcd motifs are in  
356 nucleosome-free tracts in wild-type embryos, in *bcd* mutant embryos motifs have lower  
357 overall accessibility and a graded association with nucleosomes that correlates with the  
358 sensitivity classes (41%, 46%, 50%, and 53% of motifs are accessible from high to low  
359 sensitivity). These results indicate that the mechanistic determinants of concentration-  
360 dependent Bcd action likely involve a complex interaction between Bcd, DNA, and  
361 chromatin structure.

362 **A truncated Bcd protein shows reduced binding specifically at concentration-**  
363 **sensitive target enhancers**



364 A chromatin remodeling activity associated with Bcd has not been previously  
365 described. We hypothesize that Bcd renders its target sites accessible either by  
366 competing with nucleosomes to access its binding sites and bind to DNA at high  
367 concentrations or by recruiting chromatin-remodeling enzymes to accessible motifs and  
368 subsequently driving local nucleosome remodeling to render more sites accessible. We  
369 reasoned that if Bcd can displace nucleosomes simply by competing with them for  
370 access to its binding sites, it should be possible for the Bcd DNA-binding homeodomain  
371 to compete. However, if Bcd instead drives remodeling via recruitment of cofactors, it is  
372 likely that these interactions or activities are carried out through regions of the protein  
373 outside of the DNA binding domain. To distinguish between these two possibilities, we  
374 designed a transgenic GFP-Bcd construct that is truncated downstream of the  
375 homeodomain. We modeled the truncated Bcd protein after the *bcd*<sup>085</sup> allele, which was  
376 originally classified as an "intermediate allele" of *bcd* (Frohnhöfer and Nüsslein-Volhard,  
377 1986) and reported to have weak transcriptional activating activity (Struhl et al., 1989).  
378 The truncation occurs 28 amino acids downstream of the homeodomain (Figure 5A),  
379 and the GFP-tagged protein was therefore expected to bind DNA but lack functions  
380 requiring its C-terminus. The truncated protein (known as GFP-Bcd<sup>085</sup>) forms a gradient  
381 from the anterior of the embryo, and is expressed at a similar level as a full-length GFP-  
382 Bcd (Figure 5B).

383 By ChIP-seq, we found that compared to wild-type, GFP-Bcd<sup>085</sup> binding in the  
384 Concentration-Sensitive I and II enhancer classes was significantly reduced relative to  
385 the Concentration-Sensitive II and Concentration-Insensitive classes (p-value < 0.0001  
386 in permutation test with n = 10,000 trials) (Figure 5C). Our ATAC-seq experiments

387 revealed that these classes have reduced chromatin accessibility in *bcd* mutants  
388 (Figures 3B and 4D). Taken together, these results suggest that Bcd's ability to access  
389 its concentration-sensitive targets is dependent upon activities carried out by domains in  
390 the C-terminus of the protein, likely via recruitment of a cofactor, and that its DNA  
391 binding activity alone is insufficient to drive chromatin accessibility.

### 392 **Bicoid binding sites confer anterior expression to a posterior target**

393 Overall, highly concentration-dependent targets are expressed in the anterior and  
394 are dependent on Bcd for accessibility, while less sensitive targets show more posterior  
395 expression patterns and a greater enrichment for Zld binding sites. An enhancer for  
396 *caudal* is a Concentration-Insensitive Bcd target and drives expression in the posterior  
397 of the embryo (Figure 6). This enhancer depends on Zld for chromatin accessibility, and  
398 consequently is not functional in *zld* mutants (Supplemental File 1 and Figure 6B). This  
399 supports previous findings that Zld binding contributes to allowing Bcd activation at low  
400 concentrations in posterior nuclei (Xu et al., 2014). Like the *kni* posterior enhancer  
401 (Figure 1A), the *caudal* enhancer is Bcd independent for chromatin accessibility and its  
402 expression boundary shifts anteriorly in *bcd* mutant embryos. We tested whether we  
403 could convert the properties of the *caudal* enhancer from low to high sensitivity by  
404 manipulating DNA motifs. We identified the Zld binding sites in the *caudal* enhancer  
405 sequence and mutated them to Bcd binding sites (Figure 6A). These mutations result in  
406 a shift of *caudal* reporter expression to the anterior of the embryo. Anterior expression  
407 of the mutated reporter is Bcd dependent, as it is lost in *bcd* mutant embryos.  
408 Importantly, the mutant enhancer is functional in *zld* mutant embryos, retaining a distinct  
409 anterior expression domain. In the absence of Zld, the wild type enhancer does not

410 drive expression. (Figure 6B) By replacing Zld motifs with Bcd motifs, the enhancer  
411 retains functionality, but the spatial domains of expression are now restricted to regions  
412 of high Bcd concentration. These results are consistent with a model where Bcd  
413 operates at high concentrations to confer chromatin accessibility at target sites, thereby  
414 delineating distinct gene expression and chromatin states at specific positions along its  
415 concentration gradient.

## 416 **DISCUSSION**

### 417 **A model for chromatin accessibility thresholds at Bcd target genes**

418 The results presented here demonstrate that the positional information in the Bcd  
419 gradient is read out as differential binding between Bcd and the *cis*-regulatory regions of  
420 its target genes. The overrepresentation of enhancers for anteriorly expressed target  
421 genes in the more sensitive classes provides support for the classic French flag model,  
422 as their enhancers are only capable of binding Bcd at high levels. However, motif  
423 analysis and *in vitro* EMSA experiments reveal that the differences in binding affinities  
424 that we observe *in vivo* cannot be explained entirely by the sequence of Bcd binding  
425 sites in the target enhancers. Instead we find that a subset of the enhancers in the  
426 concentration-sensitive classes require Bcd for chromatin accessibility. Taken together,  
427 this leads us to model in which the Bcd morphogen establishes concentration  
428 thresholds along the AP axis of the developing embryo by driving opening chromatin  
429 states at high concentrations, thereby gaining access to its most sensitive target  
430 enhancers. At lower concentrations in more posterior nuclei, Bcd is unable to access  
431 these enhancers, and therefore does not bind and activate their transcription. (Figure 7)  
432 In this way, expression of these concentration-sensitive target genes is restricted to

433 anterior regions of the embryo. The higher density of Bcd binding sites in highly  
434 concentration-sensitive target enhancers (shown in Figure 2B) suggests that these  
435 enhancers may require a larger number of Bcd molecules to be bound at a given time to  
436 keep the enhancers free of nucleosomes and accessible to the additional regulatory  
437 factors. We therefore provide a model for morphogen function in which the  
438 concentration thresholds in the gradient are read out molecularly at the level of  
439 chromatin accessibility, rather than through the strength of binding sites in the target  
440 enhancers.

441         It is important to note that the discrete sensitivity classes described here were  
442 generated by Bcd binding data, and this binding occurs prior to the activation of target  
443 genes and refinement of their expression domains. In our model Bcd establishes these  
444 initial patterns not by competing with its own target genes, but with default nucleosome  
445 positions in the early embryo. We predict that this initial interaction with chromatin is an  
446 essential event for establishing distinct, positionally defined patterns of gene  
447 expression. The chromatin landscapes established early by Bcd are then elaborated  
448 upon by additional patterning factors, including Bcd target genes themselves, as well as  
449 the repressor gradients proposed by Chen, *et al.* (Chen et al., 2012) Thus, the pre-  
450 transcriptional information presented by Bcd in the form of differential binding states is  
451 refined at the level of gene expression domains both by Bcd and other transcription  
452 factors active in the early embryo.

### 453 **Relationship of Bicoid and Zelda at Bicoid-bound enhancers**

454         The prominence of the Zld binding motif in the Bcd-bound ChIP peaks and  
455 ATAC-seq in *zld* mutants reveals that Zld also contributes to the accessibility of Bcd

456 targets in the genome, in part at those targets that are not dependent on Bcd for their  
457 accessibility (61/1,027 peaks are dependent on both Bcd and Zld for accessibility). Zld  
458 is therefore likely to be one component that influences the accessibility and therefore  
459 the apparent *in vivo* affinity of the enhancers that are bound by Bcd but insensitive to its  
460 local concentration (Figure 7). Previous work has suggested that Zld contributes to  
461 activation of target genes at low concentrations of Bcd protein (Xu et al., 2014). That  
462 study, in combination with the work presented here, allows us to predict that  
463 transforming a concentration-sensitive Bcd target enhancer into a Zld dependent  
464 enhancer would increase the accessibility and therefore the sensitivity of that region *in*  
465 *vivo*. Indeed, Xu, *et al.* have previously demonstrated that adding increasing numbers of  
466 Zld sites to an inactive Bcd-bound enhancer can drive increasingly posterior gene  
467 expression (Xu et al., 2014). The reporter construct used to demonstrate this effect  
468 (HC\_45) is identified as a Concentration-Sensitive I target in our study. We posit that  
469 the increase in gene expression from this reporter observed in their work is the result of  
470 increasing the accessibility of the enhancer region.

471 Alternately, when we replace Zld sites with Bcd sites in an enhancer that drives  
472 posterior expression, the expression domain shifts to the anterior of the embryo. This  
473 demonstrates that without Zld to keep the enhancer open in posterior nuclei, activation  
474 of the reporter gene becomes entirely dependent on Bcd, effectively shifting this  
475 reporter from a concentration-insensitive to concentration-sensitive enhancer. This  
476 finding fits with both previously reported findings and the model proposed in our study.  
477 Namely, that Zld contributes to the accessibility of Bcd target genes throughout the

478 embryo, while high levels of Bcd can drive accessibility independently and activate gene  
479 expression at a subset of targets in the anterior of the embryo.

480         The reduced binding by a truncated Bcd protein at the most concentration-  
481 sensitive targets indicates that Bcd does not displace nucleosomes by simply by  
482 competing for binding to genomic targets, but rather that the C-terminus of the Bcd  
483 protein is required for accessing its nucleosome-associated DNA targets. Previous work  
484 has shown that various domains of the Bcd protein are required for interactions with  
485 both co-activators and co-repressors. The N-terminus of Bcd is required for interactions  
486 with components of the Sin3A/HDAC repressor complex, and these interactions are  
487 proposed to play a role in reducing Bcd's transcriptional activation activity (Zhao et al.,  
488 2003; Zhu et al., 2001). Multiple Bcd domains, including the C-terminus, are required for  
489 interaction with CREB-binding protein (CBP), which has histone acetyltransferase  
490 activity (Fu and Ma, 2005; Fu et al., 2004). It is possible that in our truncated Bcd  
491 construct, this interaction with CBP is disrupted. As CBP is thought to play a role in  
492 increasing chromatin accessibility for transcription factors (Chan and La Thangue,  
493 2001), the loss of this interaction could lead to the reduced binding to sensitive targets  
494 that we observe in embryos with truncated Bcd. The enhancers that we have classified  
495 as concentration-insensitive do not depend on Bcd to establish an open chromatin  
496 state. This suggests that these sites are opened by other chromatin remodeling factors,  
497 or are inherently more likely to be nucleosome-free based on their underlying sequence.

498         It has previously been suggested that transcription factors can compete with  
499 nucleosomes for access to their DNA binding sites (Mirny, 2010; Wang et al., 2011).  
500 This could occur through cooperative binding to nucleosome-associated enhancers: if

501 one Bcd molecule could gain access to a binding site that was protected by a  
502 nucleosome, it could recruit additional Bcd protein molecules to bind to nearby sites and  
503 occlude nucleosome binding. This cooperativity would require a high concentration of  
504 Bcd protein, fitting with our observation that Bcd influences accessibility more strongly  
505 at high concentrations. However, our experiments with a truncated Bcd protein reveal  
506 that Bcd cannot bind to its most sensitive targets without its C-terminal domains. As  
507 many of the residues that have been implicated in cooperative binding reside in the Bcd  
508 homeodomain (Burz and Hanes, 2001), we would expect this truncated Bcd to bind  
509 cooperatively. This finding therefore supports a model in which Bcd is actively  
510 remodeling chromatin, either directly or more likely by interacting with chromatin  
511 remodeling factors through its C-terminus.

512         As a maternally supplied factor, Bcd provides one of the first cues to the break  
513 the symmetry of the embryonic body plan. Our results suggest that this symmetry  
514 breaking occurs first at the level of chromatin accessibility, as Bcd drives the opening of  
515 its most concentration-sensitive target enhancers in anterior nuclei. Another maternal  
516 factor, Zld, is proposed to act as a pioneer factor at early embryonic enhancers with a  
517 high intrinsic nucleosome barrier. By binding to these enhancers, Zld depletes them of  
518 nucleosomes and allows patterning transcription factors to bind and activate gene  
519 expression (Sun et al., 2015). We have demonstrated here that Bcd influences the  
520 accessibility primarily of its concentration-sensitive targets, which also exhibit a high  
521 predicted nucleosome barrier (Figure 4B). This raises the possibility that Bcd may be  
522 exhibiting pioneer-like activity at high concentrations, driving accessibility of these sites  
523 prior to transcriptional activation. It is unlikely that Bcd is unique in its ability to influence

524 the local chromatin accessibility of its targets. Recent work in mouse embryos has  
525 shown that another homeodomain transcription factor, Cdx2, influences the chromatin  
526 accessibility of its targets during posterior axial elongation (Amin et al., 2016). This may  
527 be a common property of developmental transcription factors that must gain early  
528 access their target enhancers while the chromatin state of the genome is being  
529 remodeled during large-scale transitions in the gene regulatory landscape.

530

## 531 **EXPERIMENTAL PROCEDURES**

### 532 **Fly stocks and Genetics**

533 *bcd* mutants refers to embryos derived from *bcd*<sup>E1</sup> homozygous mothers. The *bcd*<sup>E1</sup> and  
534 *bcd*<sup>E1</sup> *nos*<sup>L7</sup> *tsl*<sup>4</sup> stocks were from the Wieschaus/Schüpbach stock collection maintained  
535 at Princeton University. *zld* mutants are embryos derived from *zelda*<sup>294</sup> germline clones.  
536 Zelda mutant embryos were generated from the *zld*<sup>294</sup> allele (kind gift of Christine  
537 Rushlow) as germline clones as described previously (Blythe and Wieschaus, 2015).  
538 Uniform Bcd and Bcd<sup>085</sup> transgenes were expressed in a *bcd*<sup>E1</sup> mutant background.  
539 Germline clones possessing only positional information from Bicoid were generated by  
540 heat shocking *hsFLP; FRT82B hb*<sup>FB</sup> *nos*<sup>BN</sup> *tsl*<sup>4</sup> / *FRT82B tsl*<sup>4</sup> *Ovo*<sup>D</sup> larvae. Germline  
541 clones lacking Bicoid positional information as well were generated by heat shocking  
542 *hsFLP; FRT82B bcd*<sup>E1</sup> *hb*<sup>FB</sup> *nos*<sup>BN</sup> *tsl*<sup>4</sup> / *FRT82B tsl*<sup>4</sup> *Ovo*<sup>D</sup> larvae. Embryos from  
543 homozygous *eGFP-Bcd; bcd*<sup>E1</sup> *nos*<sup>L7</sup> *tsl*<sup>4</sup> mothers were used in ChIP-seq experiments to  
544 determine the impact of removing other maternal factors on Bcd binding to its targets.  
545 All ATAC-seq experiments were performed in a *His2Av-GFP* (Bloomington) background  
546 to facilitate scoring of nuclear density.



547 The uBcd and Bcd<sup>085</sup> constructs were injected for site directed transgenesis into  
548 embryos from a *y,w;atp40* stock by Genetic Services (bcd-uBcd and  $\alpha$ Tub67C-uBcd) or  
549 BestGene (mtrm-uBcd and Bcd<sup>085</sup>) and stable transformant lines were established. The  
550 mutant cad-GAL4 reporter was injected into a *M{vas-int.Dm}ZH-2A, P{CaryP}attP2*  
551 stock by Rainbow Transgenic Flies, Inc.

552 The uBcd transformants expressed eGFP-tagged Bcd in a graded distribution in the  
553 embryo and RFP in the eyes. Transgenic flies containing the uBcd constructs were  
554 crossed into a *bcd*<sup>E1</sup> background. To achieve uniform Bcd expression, the uBcd flies  
555 were crossed to a stock expressing a heat shock-inducible *flippase* in a *bcd*<sup>E1</sup>  
556 background and the resulting larvae were heat shocked at 37°C. Recombination of the  
557 FRT-flanked cassette containing the *bcd* 3'UTR and 3xP3-RFP was scored by a mosaic  
558 loss of RFP expression in the eyes. Mosaic flies were further outcrossed to *bcd*<sup>E1</sup> and  
559 progeny lacking the *bcd* 3'UTR were sorted by loss of RFP expression. The resulting  
560 flies produced embryos in which the *bcd* 3'UTR was replaced by the *sqh* 3'UTR causing  
561 a uniform distribution of Bcd along the AP axis. (Figure 1 Figure Supplement 1C)

562

### 563 **Transgenic Constructs**

564 The uniform Bcd constructs were generated using a pBABR plasmid containing an N-  
565 terminal GFP-tagged *bcd* cDNA in which the *bcd* 3'UTR was replaced by the *sqh* 3'UTR  
566 (pBABR GFP-Bcd3'sqh) (Oliver Grimm, unpublished). This results in a loss of mRNA  
567 localization at the anterior pole of the oocyte. A sequence containing the *bicoid* 3'UTR  
568 and a 3xP3-RFP reporter flanked by FRT sites was synthesized by GenScript and  
569 cloned by Gibson Assembly into the pBABR GFP-Bcd3'sqh plasmid. The FRT-flanked

570 cassette was inserted between the *bcd* coding sequence and the *sqh* 3'UTR. The *bcd*  
571 promoter was removed by digesting with AgeI and KpnI and replaced with either the  
572 *mtrm* or the *αTub67C* promoter to generate the *bcd-uBcd*, *mtrm-uBcd*, and *αTub67C-*  
573 *uBcd* constructs.

574 The GFP-Bcd<sup>085</sup> truncation was generated from eGFP-Bcd (Gregor et al., 2007a) in  
575 pBlueScript by amplifying with primers to create a stop codon after amino acid 179 as in  
576 the *bcd*<sup>085</sup> hypomorphic EMS allele (Rivera-Pomar et al., 1996). The primers inserted an  
577 *AvrII* restriction site 3' to the deletion site.

578 F Primer: 5'-TTGtagCCTAGGCCTGGATGAGAGGCGTGT-3'

579 R Primer: 5'-TCCAGGCCTAGGctaCAAGCTGGGGGGATC-3'

580 The plasmid was amplified by PCR and the linear product was digested and ligated to  
581 create the Bcd<sup>085</sup> truncation. The GFP-Bcd<sup>085</sup> construct was digested from pBlueScript  
582 with BamHI and EcoRI and ligated into pBabr.

583 The wild-type *cad*-GAL4 reporter (VT010589, coordinates chr2L: 20767347–20768825)  
584 was ordered from the Vienna Drosophila Resource Center (VDRC ID 205848/construct  
585 ID 210589). The mutated *cad* enhancer sequence was synthesized by GenScript,  
586 amplified, and cloned into the pBPGUw vector (Pfeiffer et al., 2008).

587 Primers, sequences, and plasmids are available upon request.

588

## 589 **Western Blots**

590 Live embryos were dechorionated in bleach, rinsed in salt solution (NaCl with TritonX-  
591 100), and embryos at NC14 were sorted under a light microscope and flash frozen on  
592 dry ice. Western blots were performed using a using a rabbit anti-GFP antibody

593 (Millipore Cat # AB3080P) and mouse anti-tubulin antibody (Sigma Cat # T9026) as a  
594 loading control. For quantification, the GFP band intensities were normalized to  $\alpha$ -  
595 tubulin band intensities in each lane. Two biological replicates of 50 embryos were  
596 homogenized in 50 $\mu$ L buffer for each genotype, and 10 $\mu$ L (= 10 embryos) was loaded  
597 per lane.

598 Western blots were used to generate an estimate of Bcd concentration in each of the  
599 uniform Bcd lines. Drocco, *et al.* used western blots to measure Bcd protein  
600 accumulation in the embryo during development, and calculated the total amount of Bcd  
601 in the embryo at NC14 to be  $1.5 \pm 0.2 \times 10^8$  molecules (Drocco et al., 2011). Given that the  
602 volume of the nucleus is  $\sim 1/10$  (or  $1/1+9$ ) the volume of the cytoplasm and Bcd  
603 partitions between the nucleus and the cytoplasm at a ratio of  $\sim 4:1$  (Gregor et al.,  
604 2007a), we can generate a ratio of  $4/4+9$  or 0.31 for nuclear/cytoplasmic Bcd. Using this  
605 value, we can convert  $1.5 \pm 0.2 \times 10^8$  molecules/embryo into  $4.6 \times 10^7$  molecules/nucleus at  
606 NC14. At this stage, there are 6,000 nuclei at the cortex of the embryo, which would be  
607  $\sim 7,750$  Bcd molecules/nucleus if the Bcd were distributed uniformly. Additionally, optical  
608 measurements estimate a nuclear concentration of Bcd as  $8 \pm 1$  nM and 690 Bcd  
609 molecules at the *hunchback* expression boundary ( $\sim 48\%$  x/L) at NC14 (Gregor et al.,  
610 2007b). We used these values to generate a conversion factor of 0.011594203  
611 nM/molecule and calculate the approximate nuclear concentrations given below for  
612 each uniform Bcd line. See also Figure 1 Figure Supplement 1D.

<b>Genotype</b>	<b>Expression/WT</b>	<b>Number of Molecules</b>	<b>Nuclear Concentration</b>
bcd>uBcd	0.14	1085	12.58 nM

mtrm>uBcd	1.1	8525	98.84 nM
$\alpha$ Tub67C>uBcd	2.7	20925	242.61 nM

613 Table S1. Estimated nuclear concentrations of Bcd protein in each uniform line.

#### 614 **Immunostaining and Imaging**

615 Embryos of indicated genotypes were collected from 0-4 hour laying cages, and fixed  
616 and stained essentially as described in (Dubuis et al., 2013), with rabbit anti-Bcd,  
617 guinea pig anti-Kni, and rat anti-Btd primary antibodies, followed by fluorophore-  
618 conjugated secondary antibodies Alexa-488 (guinea pig), Alexa-568 (rat), and Alexa-  
619 647 (rabbit) from Invitrogen. Stained embryos were imaged on a Leica SP5 laser-  
620 scanning confocal microscope.

621

#### 622 **Live Imaging and Image Analysis**

623 Dechorionated embryos of the indicated genotypes were mounted on coverslips  
624 overlaid with halocarbon oil and imaged in the mid-sagittal plane on a Leica SP5 laser  
625 scanning confocal microscope. Image analysis was performed in MATLAB  
626 (<http://www.mathworks.com>). GFP intensity along the dorsal profile of each embryo was  
627 extracted for each frame of the live movies in nuclear cycle 14. The frame with the  
628 highest overall intensity in each movie was plotted.

629

#### 630 **Bicoid homeodomain expression and protein purification**

631 A cDNA coding for amino acids 89-154 of the Bicoid protein (including the  
632 homeodomain) as described in (Burz et al., 1998) with a C-terminal HA epitope tag was  
633 cloned into the pET-15b plasmid, which contains an N-terminal 6xHis tag and T7

634 promoter, to make plasmid pET-15B-BcdHD. Expression was induced in BL21 (DE3)  
635 pLysS *E. coli* cells using 2 mM IPTG. The protein was purified by affinity  
636 chromatography using HisPur Cobalt Resin (Fisher Scientific Cat # 89965) followed by  
637 ion exchange chromatography with SP Sepharose Fast Flow resin (GE Healthcare Cat  
638 #17-0729-01).

### 639 EMSAs and $K_d$ Calculations

640 EMSAs were performed using purified Bicoid homeodomain and biotin-labeled DNA  
641 probes were designed to span ~200 bp in the maximal peak region of Bcd-bound peaks  
642 identified by ChIP and corresponding to previously characterized enhancers. Effective  
643  $K_d$  values for each enhancer probe were calculated using the ratio of total shifted probe  
644 to free probe.

Primers		Sequence (5'->3')
hbP2 probe F	<b>Forward primer</b>	/bio/GTCAAGGGATTAGATGGGCA
hbP2 Probe R	<b>Reverse primer</b>	/bio/GTCGACTCCTGACCAACGTA
kni post F	<b>Forward primer</b>	/bio/AGAAAAAATGAGAACAATGTGAC
kni post R	<b>Reverse primer</b>	/bio/AGCCAGCGATTTGTTACCT
kni ant F	<b>Forward primer</b>	/bio/ACAACACCGACCCGTAATCC
kni ant R	<b>Reverse primer</b>	/bio/GTCATGTTGGCTAATCTGGC
kr ant F	<b>Forward primer</b>	/bio/CAGAAAAGAAAAAGTGTAAACGCC

Kr ant R	<b>Reverse primer</b>	/bio/GCGAAAAACGCGTCGCGCT
otd intron F	<b>Forward primer</b>	/bio/ATCGTTCCTTGCGGTTTAAT
otd intron R	<b>Reverse primer</b>	/bio/AGAACAGGACAAAGGGAATTTAATC
otd early F	<b>Forward primer</b>	/bio/CTCGCCTCGCGTGCGACATT
otd early R	<b>Reverse primer</b>	/bio/CCTGCGGCAGGACTTCACTT
btd F	<b>Forward primer</b>	/bio/ACGAAGTCAAACTTTTCCA
btd R	<b>Reverse primer</b>	/bio/AGCTAAGAGATCTCAACCAAC
gt -3 F	<b>Forward primer</b>	/bio/TTACAAC TGCCCATTCAGGG
gt -3 R	<b>Reverse primer</b>	/bio/GAAGGGCTCGGGTTCGG
gt -10 F	<b>Forward primer</b>	/bio/AGATCCAGGCGAGCACTTGA
gt -10 R	<b>Reverse primer</b>	/bio/TTAAATTAAAATGTCGCAGGAAGGCG

645 Table S2: Primer sequences for EMSA probes.

646

## 647 **ChIP-seq and Data Analysis**

### 648 ***Sample Collection***

649 *Drosophila* embryos were collected from 0-4 hour laying cages, dechorionated in bleach  
650 and crosslinked in with 180 mL 20% paraformaldehyde in 2 ml PBS + 0.5% Triton X-100  
651 and 6 mL Heptane for 15 minutes. Crosslinking was quenched with 125 mM Glycine in  
652 PBS + 0.5% Triton X-100. Fixed embryos were visually staged and sorted using a  
653 dissecting microscope, and all experimental replicates consisted of 200 embryos in

654 nuclear cycle 14. Chromatin immunoprecipitation was performed with an anti-GFP  
655 antibody (Millipore) in embryos expressing GFP-tagged Bcd either in a wild-type graded  
656 distribution (*eGFP-Bcd;;bcd<sup>E1</sup>* and *eGFP-Bcd;;bcd<sup>E1</sup> nos<sup>L7</sup> tsl<sup>A</sup>*) or uniformly (*GFP-*  
657 *uBcd;;bcd<sup>E1</sup>*). Sequencing libraries were prepared using the NEBNext ChIP-seq Library  
658 Prep master mix kit and sequenced as described in (Blythe and Wieschaus, 2015;  
659 Drocco et al., 2011).

### 660 ***Defining a Peak List***

661 Barcode split sequencing files were mapped to *Drosophila melanogaster* genome  
662 assembly BDGP R5/dm3 using Bowtie2 (Gregor et al., 2007a; Langmead and Salzberg,  
663 2012) using default parameters. To generate a conservative, high-confidence list of  
664 Bcd-bound peaks, peaks were called on each replicate of wild-type and uniform Bcd  
665 ChIP-seq data using MACS2 (Gregor et al., 2007b; Zhang et al., 2008) with settings -p  
666 1e-3 --to-large --nomodel --shiftsize 130 for wild-type samples and -p 0.000001 --slocal  
667 5000 --llocal 50000 --keep-dup all for uBcd samples. The most reproducible peaks from  
668 each genotype were selected using an irreproducible discovery rate (IDR) of 1%  
669 (Dubuis et al., 2013; Landt et al., 2012; Li et al., 2011). Given evidence that highly  
670 transcribed (i.e., highly accessible) regions often give false positive results in ChIP  
671 experiments (Burz et al., 1998; Teytelman et al., 2013), we used our ATAC-seq data to  
672 filter our ChIP-seq peaks. We compared the number of CPM-normalized ATAC-seq  
673 reads to ChIP-seq reads in each peak, and performed permutation tests (n = 1,000) to  
674 determine the probability of selecting open regions of the genome at random that had  
675 higher ATAC-seq counts (i.e., regions that were more accessible) than the ATAC-seq  
676 counts in the Bcd ChIP-peaks. We determined that at a ratio of 5.4 ATAC-seq/ChIP-seq

677 counts, 95% of the ChIP peaks (permutation test p value = 0.05) were no more open  
678 than a random selection of open regions. We filtered out the remaining ChIP peaks with  
679 ATAC/ChIP ratios above 5.4, as these peaks are more likely to correspond to highly  
680 transcribed open regions where most false positive signals can be found. We then  
681 chose the peaks that were common to wild-type and uniform Bcd embryos, which  
682 resulted in a list of 1,027 Bcd-bound peak regions. The number of peaks at each step of  
683 this filtering is shown in Table S3.

Filter Applied	Number of Peaks			
	Wild-Type	tub>uBcd	mtrm>uBcd	bcd>uBcd
<b>MACS2</b>	29,090	15,429	11,812	38,392
<b>IDR</b>	9,815	4,245	1,464	1,329
<b>Euchromatic only</b>	2,319	4,123	1,429	1,257
<b>Common peaks (2/3)</b>		4,126		
<b>ATAC-seq ratios</b>	2,143	2,087		
<b>Common Peaks</b>	1,027			

684 Table S3. Number Bcd ChIP-seq peaks at each step of filtering.

685

### 686 ***Comparing Binding Between Uniform Bcd Levels***

687 Mapped BAM files were imported into R as GenomicRanges objects (Lawrence et al.,  
688 2013), filtering out reads with map quality scores below 30. Significant differences  
689 between the uBcd levels were assessed on a pairwise basis using edgeR (Robinson et  
690 al., 2010) in the 1,027 pre-defined peak plus 50,000 additional non-peak noise regions  
691 selected from the dm3 genome.



## 692 **Data Normalization and Display**

693 Sequencing data was z-score normalized for display in heatmaps. Sequencing read  
694 count coverage was calculated for 10 base pair windows across the genome, and the  
695 mean counts per million reads were determined in each ChIP peak, as well as the  
696 additional noise peaks. Z-scores were computed for each peak using

$$697 \quad z = \frac{CPM - \mu}{\sigma}$$

698 where  $\mu$  = mean CPM in noise peaks and  $\sigma$  = standard deviation of CPM in noise  
699 peaks.

## 700 **Overlaps with Vienna Tile-GAL4 Enhancers**

	Overlaps		Total Vienna Tiles
	Bcd Peaks	Vienna Tiles	
<b>Expressed (all stages)</b>	193	238	3604
<b>Expressed (stage 4-6)</b>	151	163	666
<b>Not Expressed</b>	41	55	4189
<b>Total</b>	234	293	7793

701 Table S4. Number of overlapping Bcd ChIP peaks and Vienna Tile-GAL4 enhancer  
702 reporters. The reporters expressed at stage 4-6 that overlapped with more than one Bcd  
703 peak were excluded from the plot in Figure 2B.  
704

## 705 **ATAC-seq and Data Analysis**

### 706 **Sample Collection**

707 Live embryos expressing a histone (H2Av)-GFP or RFP construct were individually  
708 staged on an epifluorescence microscope in halocarbon oil. After the onset of nuclear  
709 cycle 14, single embryos were dechorionated in bleach and macerated in cold lysis  
710 buffer at  $t = 12$  minutes into NC14. Samples were pelleted in lysis buffer at  $4^{\circ}\text{C}$  (3000  
711 rpm for 10 minutes), buffer was removed, and the embryo pellet was flash frozen on dry

712 ice. Frozen pellets were resuspended in Nextera Tagment DNA Buffer + Enzyme and  
713 incubated at 37°C for 30 minutes shaking at 800 rpm. Tagged DNA was purified using a  
714 Qiagen Minelute column and eluted in 10 µL. Barcoded sequencing libraries were  
715 generated by PCR amplifying the purified DNA using the Nextera DNA Sample Prep kit.  
716 Paired-end sequencing was performed on six samples per genotype by the Lewis Sigler  
717 Institute for Integrative Genomics Sequencing Core Facility on an Illumina HiSeq 2500.

### 718 ***Data Processing***

719 Initial processing of the data was performed essentially as described in (Blythe and  
720 Wieschhaus 2016, *submitted*). Sequencing files were barcode split and adaptors were  
721 trimmed using TrimGalore. Trimmed reads were mapped to the BDGP R5/dm3 genome  
722 assembly using BWA (Li and Durbin, 2009) with default parameters. Optical and PCR  
723 duplicates were marked using Picard Tools MarkDuplicates  
724 (<https://broadinstitute.github.io/picard/>). Mapped reads were filtered using samtools (Li  
725 et al., 2009) to remove reads with a map quality score  $\leq 30$ , unmapped reads,  
726 improperly paired reads, and duplicate reads. To distinguish reads corresponding to  
727 open chromatin reads from nucleosome protected reads, the size of the ATAC-seq  
728 fragments were fit to the sum of an exponential and Gaussian distribution as described  
729 in (Buenrostro et al., 2013). We used a fragment size cutoff of  $\leq 100$  bp to identify  
730 fragments originating from open chromatin. Filtered open reads were imported into R as  
731 GenomicRanges objects.

### 732 ***Peak Calling***

733 Regions of open chromatin at NC14+12 minutes were determined by calling peaks on  
734 the merged open chromatin reads from wild-type replicates using Zinba (Rashid et al.,  
735 2011) with the parameters:

736 input = 'none', winSize = 300, offset = 50, extension = 65, selectmodel = FALSE,  
737 formula = exp\_count ~ exp\_cnvwin\_log + align\_perc, formulaE = exp\_count ~  
738 exp\_cnvwin\_log + align\_perc, formulaZ = exp\_count ~ align\_perc, FDR = TRUE,  
739 threshold = 0.05, winGap = 0, cnvWinSize = 2.5E+4, refinepeaks = TRUE.

#### 740 ***Nucleosome Positioning***

741 Nucleosome positioning was determined samples from all genotypes using  
742 NucleoATAC (Schep et al., 2015), with default settings. Peak regions used for  
743 NucleoATAC were open chromatin peaks from Zinba combined with the Bcd CHIP  
744 peaks and widened to 2500 bp centered over the peak maxima. This combined peak list  
745 was then reduced and used as NucleoATAC input. For each genotype, BAM files from  
746 ATAC-seq were merged and used as input for NucleoATAC.

#### 747 ***Differential Accessibility between Wild-Type, $bcd^{E1}$ , $zld$ Embryos***

748 EdgeR was used to determine significant differences in accessibility between different  
749 genotypes, with an exact test FDR  $\leq 0.05$  used as the significance cutoff. For edgeR  
750 comparisons, the ATAC-seq peaks called by Zinba from wild-type embryos at NC14+12  
751 minutes were combined with Bcd CHIP-seq peaks resized to 300 bp centered around  
752 the peak summit and 25,000 background regions. The background regions were  
753 generated by extending each of the ATAC open peaks to 10 kb and subtracting them  
754 from the dm3 genome assembly. The remaining non-peak regions were then sampled  
755 randomly 25,000 times and widened to reflect the distribution of sizes in the ATAC-seq

756 peaks. Bcd- or Zld-dependent peaks were those peaks identified as having reduced  
757 accessibility *bcd*<sup>E1</sup> or *zld* embryos. A summary of the differential accessibility in the  
758 ATAC-seq vs. Bcd ChIP-seq peaks is shown in Table S5.

<b>Peak List</b>	<b>Bcd Dependent</b>	<b>Zld Dependent</b>	<b>Bcd+Zld Dependent</b>	<b>Total</b>
ATAC Open Peaks	326	2,675	206	13,226
Bcd ChIP Peaks	132	402	61	1,027
ATAC+ChIP Common Peaks	121	379	58	855

759 Table S5. Number of peaks dependent on Bcd or Zld for chromatin accessibility. ATAC  
760 + ChIP common peaks are peaks that overlap between the Bcd ChIP-seq peaks and  
761 the wild-type ATAC-seq open chromatin peaks.  
762

763 To measure differential accessibility of Bcd motifs between wild type and *bcdE1* mutant  
764 embryos, the positions of Bcd motifs within ChIP-seq peaks were found, and ATAC-seq  
765 accessibility scores were calculated for the 10 bp window containing the midpoint of  
766 each Bcd motif. Scores for all motifs within a single peak were averaged prior to  
767 plotting. Motifs were found using the “strong” Bcd position weight matrix via the R  
768 function ‘matchPWM’ in the Biostrings package with an 80% match threshold.

769  
770 To estimate the overlap between nucleosomes and Bcd motifs, predicted nucleosome  
771 dyad centers from NucleoATAC were widened to 160 bp and motifs overlapping these  
772 intervals were scored as ‘nucleosome associated’. Motifs not overlapping widened  
773 nucleosome intervals were scored as ‘open’. The fraction of open Bcd motifs per peak  
774 was calculated by dividing the number of open Bcd motifs by the total number of  
775 encoded Bcd motifs over each peak.

776

777 ***Data Normalization and Display***

778 For each genotype, BAM-formatted ATAC-seq reads for each replicate, filtered by  
779 quality and duplicates removed as described above, were merged into a single file.  
780 Coverage was calculated from these BAM files in 10 bp windows in the dm3 genome  
781 assembly. This coverage was then normalized by counts per million reads. The genome  
782 coverage in 2 kb regions flanking Bcd ChIP peaks was then z-score normalized as  
783 described above. The z-score values are displayed as heatmaps of open chromatin  
784 displayed in Figure 3 A and B. Heatmaps are plotted in order of decreasing accessibility  
785 z score in wild-type embryos.

786 Predicted dyad centers from NucleoATAC were widened to 147 bp to model  
787 nucleosome positions. Occupancy scores from NucleoATAC in the 147 bp  
788 nucleosomes were computed in 10 bp windows across the dm3 genome assembly.  
789 Occupancy scores overlapping Bcd ChIP peaks are plotted in the heatmaps in Fig 4 A  
790 and B. Nucleosome heatmaps are plotted in the same order as open chromatin  
791 heatmaps, by decreasing accessibility z score in wild-type embryos.

792

793 Supplemental File 1 shows the genomic (BDGP Release 5/dm3) coordinates of Bcd-  
794 bound peaks identified by ChIP-seq as described in Table S3. Additional columns  
795 indicate the nearest gene to the peak that shows maternal or zygotic expression in the  
796 embryo (Blythe and Wieschaus, 2015) and classifications of each peak as Bcd or Zld  
797 dependent for accessibility (determined by ATAC-seq) and their sensitivity group  
798 classifications (determined by ChIP-seq). The tileID column gives the name of each

799 Vienna Tile-GAL4 construct (if any) that overlaps with each peak, and the HC\_ID  
800 column indicates reporter constructs from (Chen, et al., 2012) that overlap with each  
801 peak.

802

### 803 ***In situ* hybridizations and Cad-GAL4 reporters**

804 The GAL4 coding sequence was amplified from a genomic DNA preparation generated  
805 from a *Drosophila* stock carrying a GAL4 reporter, using primers 5'-

806 TGCGATATTTGCCGACTTA-3' and 5'-

807 TGTAATACGACTCACTATAGGGAACATCCCTGTAGTGATTCCA-3'. The amplified

808 sequence was used as a template in the MEGAscript T7 Transcription Kit

809 (ThermoFisher Cat. #AM1334) to with digoxigenin-labeled UTP to generate an anti-

810 GAL4 RNA probe. *In situ* hybridizations were performed according to standard

811 protocols.

812

### 813 **Author Contributions**

814 C.E.H and S.A.B. designed experiments. C.E.H. performed all experiments and C.E.H.

815 and S.A.B. performed data analysis. E.F.W. discussed results and performed difficult

816 meiotic recombination. C.E.H. wrote the first draft of the manuscript and S.A.B. and

817 E.F.W. contributed to the final version.

818

### 819 **Acknowledgements**

820 We thank T. Gregor, E. Gavis, O. Grimm, and members of the Wieschaus and

821 Schüpbach labs for helpful feedback and discussion, H. Chen, M. Levine and S. Little

822 for comments on the manuscript, and W. Wang and the staff of the Lewis-Sigler Institute  
823 Sequencing Core Facility. This work was supported by Ruth L. Kirschstein NRSA pre-  
824 doctoral fellowship F31HD082940 (C.E.H.) and postdoctoral fellowship F32HD072653  
825 (S.A.B.).  
826

## 827 **References**

- 828 Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S.C., Kester, L., van  
829 Oudenaarden, A., Creyghton, M.P., and Deschamps, J. (2016). Cdx and T Brachyury  
830 Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Rep* *17*,  
831 3165–3177.
- 832 Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., and  
833 Nüsslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the  
834 anterior pattern of the *Drosophila* embryo. *Embo J.* *7*, 1749–1756.
- 835 Blythe, S.A., and Wieschaus, E.F. (2015). Zygotic Genome Activation Triggers the DNA  
836 Replication Checkpoint at the Midblastula Transition. *Cell* *160*, 1–13.
- 837 Bosch, ten, J.R., Benavides, J.A., and Cline, T.W. (2006). The TAGteam DNA motif  
838 controls the timing of *Drosophila* pre-blastoderm transcription. *Development* *133*, 1967–  
839 1977.
- 840 Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013).  
841 Transposition of native chromatin for fast and sensitive epigenomic profiling of open  
842 chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* *10*, 1213–  
843 1218.
- 844 Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A  
845 Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* *109*,  
846 21.29.1–.29.9.
- 847 Burz, D.S., and Hanes, S.D. (2001). Isolation of mutations that disrupt cooperative DNA  
848 binding by the *Drosophila* bicoid protein. *J. Mol. Biol.* *305*, 219–230.
- 849 Burz, D.S., Rivera-Pomar, R., Jäckle, H., and Hanes, S.D. (1998). Cooperative DNA-  
850 binding by Bicoid provides a mechanism for threshold-dependent gene activation in the  
851 *Drosophila* embryo. *Embo J.* *17*, 5998–6009.
- 852 Chan, H.M., and La Thangue, N.B. (2001). p300/CBP proteins: HATs for transcriptional  
853 bridges and scaffolds. *J. Cell. Sci.* *114*, 2363–2373.
- 854 Chen, H., Xu, Z., Mei, C., Yu, D., and Small, S. (2012). A System of Repressor  
855 Gradients Spatially Organizes the Boundaries of Bicoid-Dependent Target Genes. *Cell*  
856 *149*, 618–629.
- 857 Crauk, O., and Dostatni, N. (2005). Bicoid Determines Sharp and Precise Target Gene  
858 Expression in the *Drosophila* Embryo. *Current Biology* *15*, 1888–1898.
- 859 De Renzis, S., Elemento, O., Tavazoie, S., and Wieschaus, E.F. (2007). Unmasking  
860 Activation of the Zygotic Genome Using Chromosomal Deletions in the *Drosophila*  
861 Embryo. *PLoS Biol.* *5*, e117.



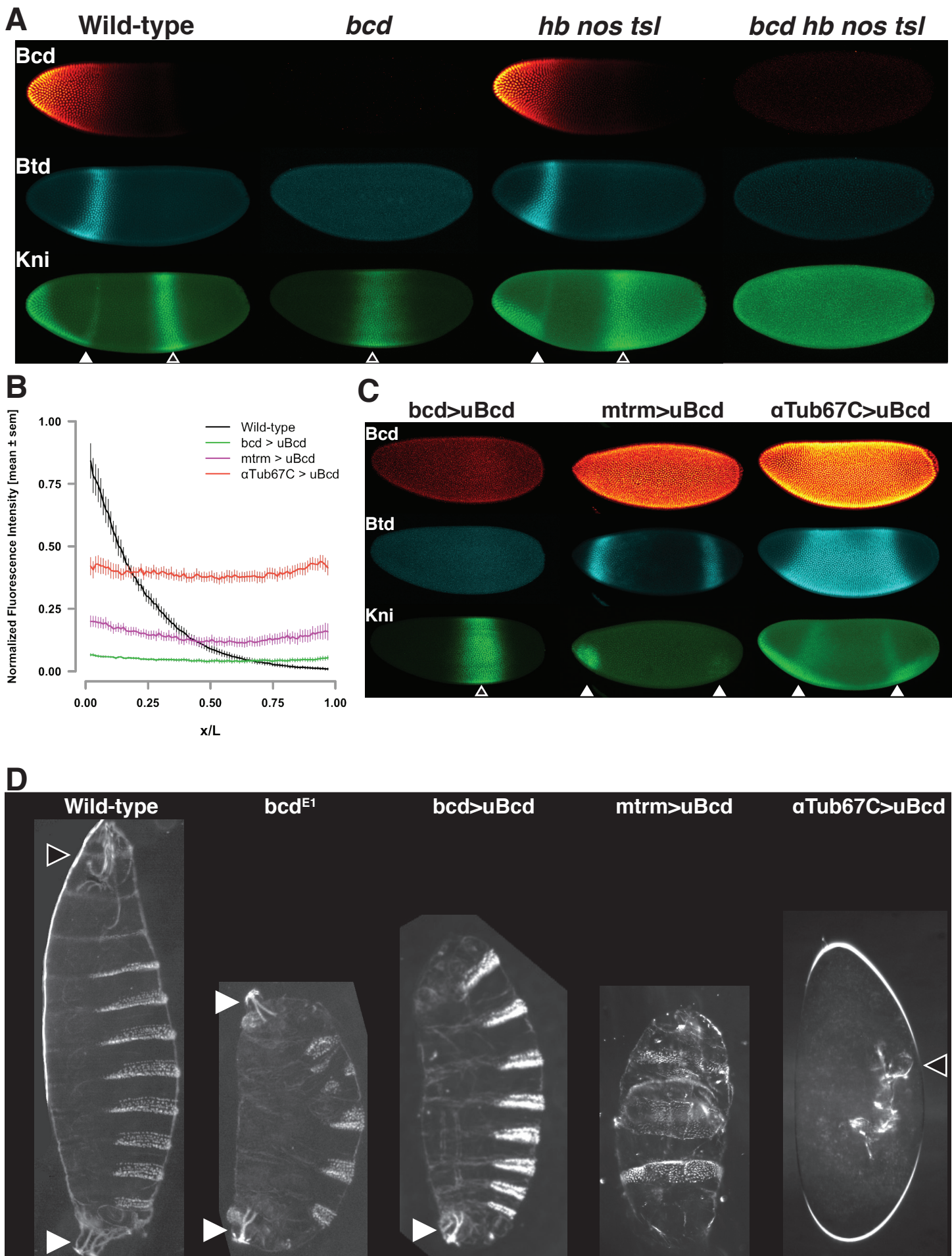
- 862 Driever, W., and Nüsslein-Volhard, C. (1988a). The bicoid protein determines position in  
863 the *Drosophila* embryo in a concentration-dependent manner. *Cell* 54, 95–104.
- 864 Driever, W., and Nüsslein-Volhard, C. (1989). The bicoid protein is a positive regulator  
865 of hunchback transcription in the early *Drosophila* embryo. *Nature* 337, 138–143.
- 866 Driever, W., Ma, J., Nüsslein-Volhard, C., and Ptashne, M. (1989a). Rescue of bicoid  
867 mutant *Drosophila* embryos by bicoid fusion proteins containing heterologous activating  
868 sequences. *Nature* 342, 149–154.
- 869 Driever, W., and Nüsslein-Volhard, C. (1988b). A gradient of bicoid protein in *Drosophila*  
870 embryos. *Cell* 54, 83–93.
- 871 Driever, W., Thoma, G., and Nüsslein-Volhard, C. (1989b). Determination of spatial  
872 domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding  
873 sites for the bicoid morphogen. *Nature* 340, 363–367.
- 874 Drocco, J.A., Grimm, O., Tank, D.W., and Wieschaus, E. (2011). Measurement and  
875 Perturbation of Morphogen Lifetime: Effects on Gradient Shape. *Biophysical Journal*  
876 101, 1807–1815.
- 877 Dubuis, J.O., Samanta, R., and Gregor, T. (2013). Accurate measurements of dynamics  
878 and reproducibility in small genetic networks. *Molecular Systems Biology* 9, 1–15.
- 879 Frohnhöfer, H.G., and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in  
880 the *Drosophila* embryo by the maternal gene bicoid. *Nature* 324, 120–125.
- 881 Fu, D., and Ma, J. (2005). Interplay between positive and negative activities that  
882 influence the role of Bicoid in transcription. *Nucleic Acids Res.* 33, 3985–3993.
- 883 Fu, D., Wen, Y., and Ma, J. (2004). The co-activator CREB-binding protein participates  
884 in enhancer-dependent activities of bicoid. *J. Biol. Chem.* 279, 48725–48733.
- 885 Gao, Q., and Finkelstein, R. (1998). Targeting gene expression to the head: the  
886 *Drosophila* orthodenticle gene is a direct target of the Bicoid morphogen. *Development*  
887 125, 4185–4193.
- 888 Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of  
889 even skipped expression are controlled by distinct regulatory elements that respond to  
890 different spatial cues. *Cell* 57, 413–422.
- 891 Gregor, T., Wieschaus, E.F., McGregor, A.P., and Bialek, W. (2007a). Stability and  
892 Nuclear Dynamics of the Bicoid Morphogen Gradient. *Cell* 130, 141–152.
- 893 Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. (2007b). Probing the limits to  
894 positional information. *Cell* 130, 153–164.
- 895 Harrison, M.M., Li, X.-Y., Kaplan, T., Botchan, M.R., and Eisen, M.B. (2011). Zelda

- 896 binding in the early *Drosophila melanogaster* embryo marks regions subsequently  
897 activated at the maternal-to-zygotic transition. *PLoS Genet.* 7, e1002266.
- 898 Häder, T., Wainwright, D., Shandala, T., Saint, R., Taubert, H., Brönner, G., and Jäckle,  
899 H. (2000). Receptor tyrosine kinase signaling regulates different modes of Groucho-  
900 dependent control of Dorsal. *Curr. Biol.* 10, 51–54.
- 901 Hoch, M., Seifert, E., and Jäckle, H. (1991). Gene expression mediated by cis-acting  
902 sequences of the Krüppel gene in response to the *Drosophila* morphogens bicoid and  
903 hunchback. *Embo J.* 10, 2267–2278.
- 904 Jaeger, J. (2010). The gap gene network. *Cell. Mol. Life Sci.* 68, 243–274.
- 905 Janody, F., Sturny, R., Catala, F., Desplan, C., and Dostatni, N. (2000). Phosphorylation  
906 of bicoid on MAP-kinase sites: contribution to its interaction with the torso pathway.  
907 *Development* 127, 279–289.
- 908 Janody, F., Sturny, R., Schaeffer, V., Azou, Y., and Dostatni, N. (2001). Two distinct  
909 domains of Bicoid mediate its transcriptional downregulation by the Torso pathway.  
910 *Development* 128, 2281–2290.
- 911 Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K.,  
912 Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of  
913 *Drosophila* developmental enhancers in vivo. *Nature*.
- 914 Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S.,  
915 Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., et al. (2012). ChIP-seq guidelines  
916 and practices of the ENCODE and modENCODE consortia. *Genome Res.* 22, 1813–  
917 1831.
- 918 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2.  
919 *Nat. Methods* 9, 357–359.
- 920 Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan,  
921 M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges.  
922 *PLoS Comput. Biol.* 9, e1003118.
- 923 Levine, M., and Davidson, E.H. (2005). Gene regulatory networks for development.  
924 *Proceedings of the National Academy of Sciences* 102, 4936–4942.
- 925 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,  
926 Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup (2009). The  
927 Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- 928 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-  
929 Wheeler transform. *Bioinformatics* 25, 1754–1760.
- 930 Li, Q., Brown, J.B., Huang, H., and Bickel, P.J. (2011). Measuring Reproducibility of

- 931 High-Throughput Experiments. *The Annals of Applied Statistics* 5, 1752–1779.
- 932 Li, X.-Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D.A., Iyer, V.N., Hechmer, A.,  
933 Simirenko, L., Stapleton, M., Luengo Hendriks, C.L., et al. (2008). Transcription factors  
934 bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.*  
935 6, e27.
- 936 Liu, F., Morrison, A.H., and Gregor, T. (2013). Dynamic interpretation of maternal inputs  
937 by the *Drosophila* segmentation gene network. *Proc Natl Acad Sci U S A* 110, 6724–  
938 6729.
- 939 Löhr, U., Chung, H.-R., Beller, M., and Jäckle, H. (2009). Antagonistic action of Bicoid  
940 and the repressor Capicua determines the spatial limits of *Drosophila* head gene  
941 expression domains. *Proceedings of the National Academy of Sciences* 106, 21695–  
942 21700.
- 943 Ma, X. (1996). Sequences Outside the Homeodomain of Bicoid Are Required for  
944 Protein-Protein Interaction. *Journal of Biological Chemistry* 271, 21660–21665.
- 945 Ma, X., Yuan, D., Diepold, K., Scarborough, T., and Ma, J. (1996). The *Drosophila*  
946 morphogenetic protein Bicoid binds DNA cooperatively. *Development* 122, 1195–1206.
- 947 MacArthur, S., Li, X.-Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondona, B.P.,  
948 Hechmer, A., Simirenko, L., Keränen, S.V., et al. (2009). Developmental roles of 21  
949 *Drosophila* transcription factors are determined by quantitative differences in binding to  
950 an overlapping set of thousands of genomic regions. *Genome Biol.* 10, R80.
- 951 Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors.  
952 *Proceedings of the National Academy of Sciences* 107, 22534–22539.
- 953 Nien, C.-Y., Liang, H.-L., Butcher, S., Sun, Y., Fu, S., Gocha, T., Kirov, N., Manak, J.R.,  
954 and Rushlow, C. (2011). Temporal coordination of gene networks by Zelda in the early  
955 *Drosophila* embryo. *PLoS Genet.* 7, e1002339.
- 956 Ochoa-Espinosa, A., Yu, D., Tsigos, A., Struffi, P., and Small, S. (2009). Anterior-  
957 posterior positional information in the absence of a strong Bicoid gradient. *Proceedings*  
958 *of the National Academy of Sciences* 106, 3823–3828.
- 959 Ochoa-Espinosa, A., Yucel, G., Kaplan, L., Pare, A., Pura, N., Oberstein, A.,  
960 Papatsenko, D., and Small, S. (2005). The role of binding site cluster strength in Bicoid-  
961 dependent patterning in *Drosophila*. *Proc Natl Acad Sci U S A* 102, 4960–4965.
- 962 Pankratz, M.J., Busch, M., Hoch, M., Seifert, E., and Jäckle, H. (1992). Spatial control of  
963 the gap gene knirps in the *Drosophila* embryo by posterior morphogen system. *Science*  
964 255, 986–989.
- 965 Perry, M.W., Boettiger, A.N., and Levine, M. (2011). Multiple enhancers ensure  
966 precision of gap gene-expression patterns in the *Drosophila* embryo. *Proc Natl Acad Sci*

- 967 U S A 108, 13570–13575.
- 968 Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.-T.B., Misra, S., Murphy, C., Scully,  
969 A., Carlson, J.W., Wan, K.H., Lavery, T.R., et al. (2008). Tools for neuroanatomy and  
970 neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences* 105,  
971 9715–9720.
- 972 Rashid, N.U., Giresi, P.G., Ibrahim, J.G., Sun, W., and Lieb, J.D. (2011). ZINBA  
973 integrates local covariates with DNA-seq data to identify broad and narrow regions of  
974 enrichment, even within amplified genomic regions. *Genome Biol.* 12, R67.
- 975 Rechsteiner, M., and Rogers, S.W. (1996). PEST sequences and regulation by  
976 proteolysis. *Trends Biochem. Sci.* 21, 267–271.
- 977 Rivera-Pomar, R., Niessing, D., Schmidt-Ott, U., Gehring, W.J., and Jäckle, H. (1996).  
978 RNA binding and translational suppression by bicoid. *Nature* 379, 746–749.
- 979 Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor  
980 package for differential expression analysis of digital gene expression data.  
981 *Bioinformatics* 26, 139–140.
- 982 Sauer, F., Hansen, S.K., and Tjian, R. (1995). DNA Template and Activator-Coactivator  
983 Requirements for Transcriptional Synergism by *Drosophila* Bicoid. *Science* 270, 1825–  
984 1828.
- 985 Schep, A.N., Buenrostro, J.D., Denny, S.K., Schwartz, K., Sherlock, G., and Greenleaf,  
986 W.J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of  
987 chromatin architecture within regulatory regions. *Genome Res.* 25, 1757–1770.
- 988 Schroeder, M.D., Pearce, M., Fak, J., Fan, H., Unnerstall, U., Emberly, E., Rajewsky,  
989 N., Siggia, E.D., and Gaul, U. (2004). Transcriptional control in the segmentation gene  
990 network of *Drosophila*. *PLoS Biol.* 2, E271.
- 991 Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang,  
992 J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* 442,  
993 772–778.
- 994 Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., and Gaul, U. (2008).  
995 Predicting expression patterns from regulatory sequence in *Drosophila* segmentation.  
996 *Nature* 451, 535–540.
- 997 Struhl, G., Struhl, K., and Macdonald, P.M. (1989). The gradient morphogen bicoid is a  
998 concentration-dependent transcriptional activator. *Cell* 57, 1259–1273.
- 999 Sun, Y., Nien, C.-Y., Chen, K., Liu, H.-Y., Johnston, J., Zeitlinger, J., and Rushlow, C.  
1000 (2015). Zelda overcomes the high intrinsic nucleosome barrier at enhancers during  
1001 *Drosophila* zygotic genome activation. *Genome Res.* 25, 1703–1714.

- 1002 Teytelman, L., Thurtle, D.M., Rine, J., and van Oudenaarden, A. (2013). Highly  
1003 expressed loci are vulnerable to misleading ChIP localization of multiple unrelated  
1004 proteins. *Proceedings of the National Academy of Sciences* *110*, 18602–18607.
- 1005 Thomas-Chollier, M., Herrmann, C., Defrance, M., Sand, O., Thieffry, D., and van  
1006 Helden, J. (2012). RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets.  
1007 *Nucleic Acids Res.* *40*, e31.
- 1008 Thomas-Chollier, M., Sand, O., Turatsinze, J.-V., Janky, R., Defrance, M., Vervisch, E.,  
1009 Brohée, S., and van Helden, J. (2008). RSAT: regulatory sequence analysis tools.  
1010 *Nucleic Acids Res.* *36*, W119–W127.
- 1011 Wang, X., Bai, L., Bryant, G.O., and Ptashne, M. (2011). Nucleosomes and the  
1012 accessibility problem. *Trends Genet.* *27*, 487–492.
- 1013 Wimmer, E.A., Simpson-Brose, M., Cohen, S.M., Desplan, C., and Jäckle, H. (1995).  
1014 Trans- and cis-acting requirements for blastodermal expression of the head gap gene  
1015 buttonhead. *Mech. Dev.* *53*, 235–245.
- 1016 Wolpert, L. (1969). Positional information and the spatial pattern of cellular  
1017 differentiation. *Journal of Theoretical Biology* *25*, 1–47.
- 1018 Xi, L., Fondufe-Mittendorf, Y., Xia, L., Flatow, J., Widom, J., and Wang, J.-P. (2010).  
1019 Predicting nucleosome positioning using a duration Hidden Markov Model. *BMC*  
1020 *Bioinformatics* *11*, 346.
- 1021 Xu, Z., Chen, H., Ling, J., Yu, D., Struffi, P., and Small, S. (2014). Impacts of the  
1022 ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in  
1023 *Drosophila*. *Genes Dev.* *28*, 608–621.
- 1024 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E.,  
1025 Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of  
1026 ChIP-Seq (MACS). *Genome Biol.* *9*, R137.
- 1027 Zhao, C., Fu, D., Dave, V., and Ma, J. (2003). A composite motif of the *Drosophila*  
1028 morphogenetic protein bicoid critical to transcription control. *J. Biol. Chem.* *278*, 43901–  
1029 43909.
- 1030 Zhu, W., Foehr, M., Jaynes, J.B., and Hanes, S.D. (2001). *Drosophila* SAP18, a  
1031 member of the Sin3/Rpd3 histone deacetylase complex, interacts with Bicoid and  
1032 inhibits its activity. *Development Genes and Evolution* *211*, 109–117.
- 1033
- 1034



**Figure 1. Uniform Bcd expression specifies cell fates corresponding to levels of expression.**

(A) Wild-type, *bcd* null mutant (*bcd*<sup>E1</sup>), and maternal *hunchback*, *nanos*, *torso-like* (*hb nos tsl*) triple mutant and *bcd hb nos tsl* mutant embryos at NC14 immunostained with antibodies against Bcd, Btd, and Kni. Embryos are oriented with anterior to the left. The anterior Kni domain (filled arrow) is absent in *bcd* but restored in *hb nos tsl* embryos, while the posterior stripe (open arrow) shifts anteriorly in *bcd* but expands posteriorly in *hb nos tsl*. Neither Btd nor Kni exhibit patterned expression in *bcd hb nos tsl*. Images are maximum z-projections and image contrast was adjusted uniformly across the entire image for display. See Figure S1A for quantification of Kni intensity between genotypes.

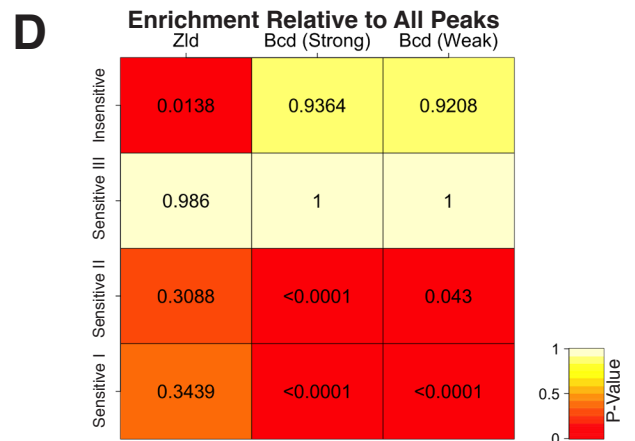
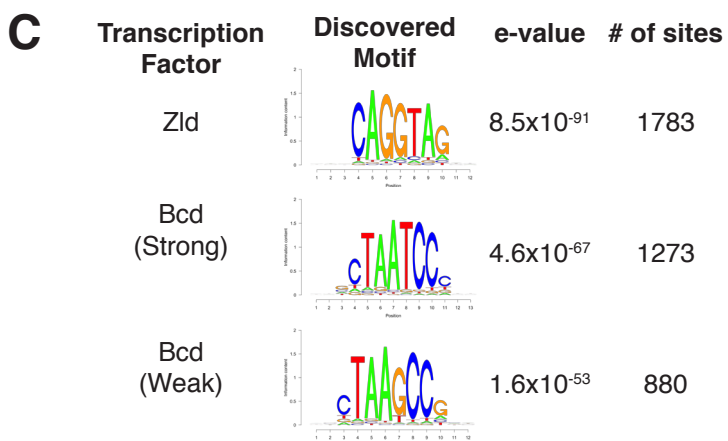
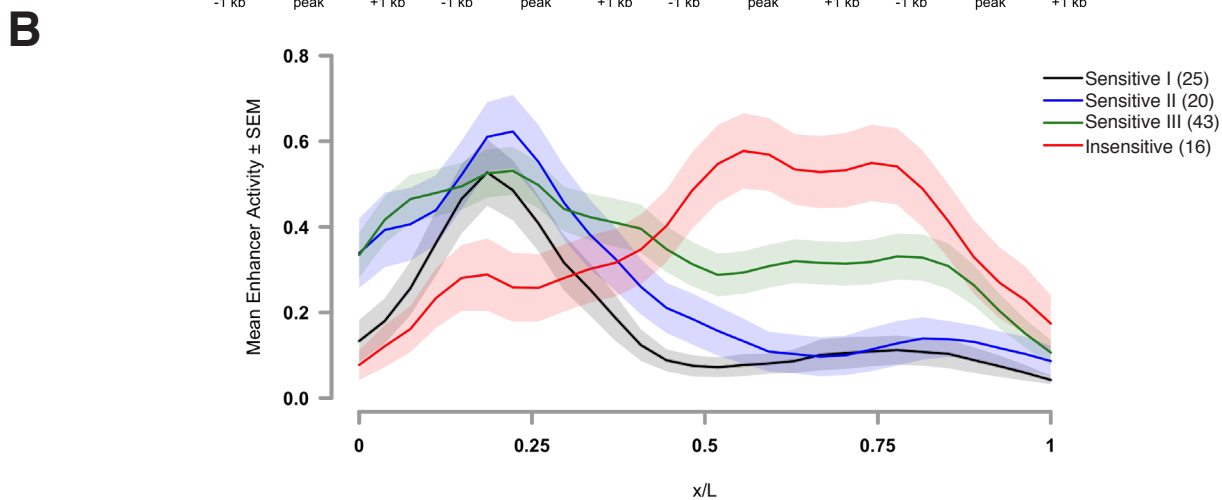
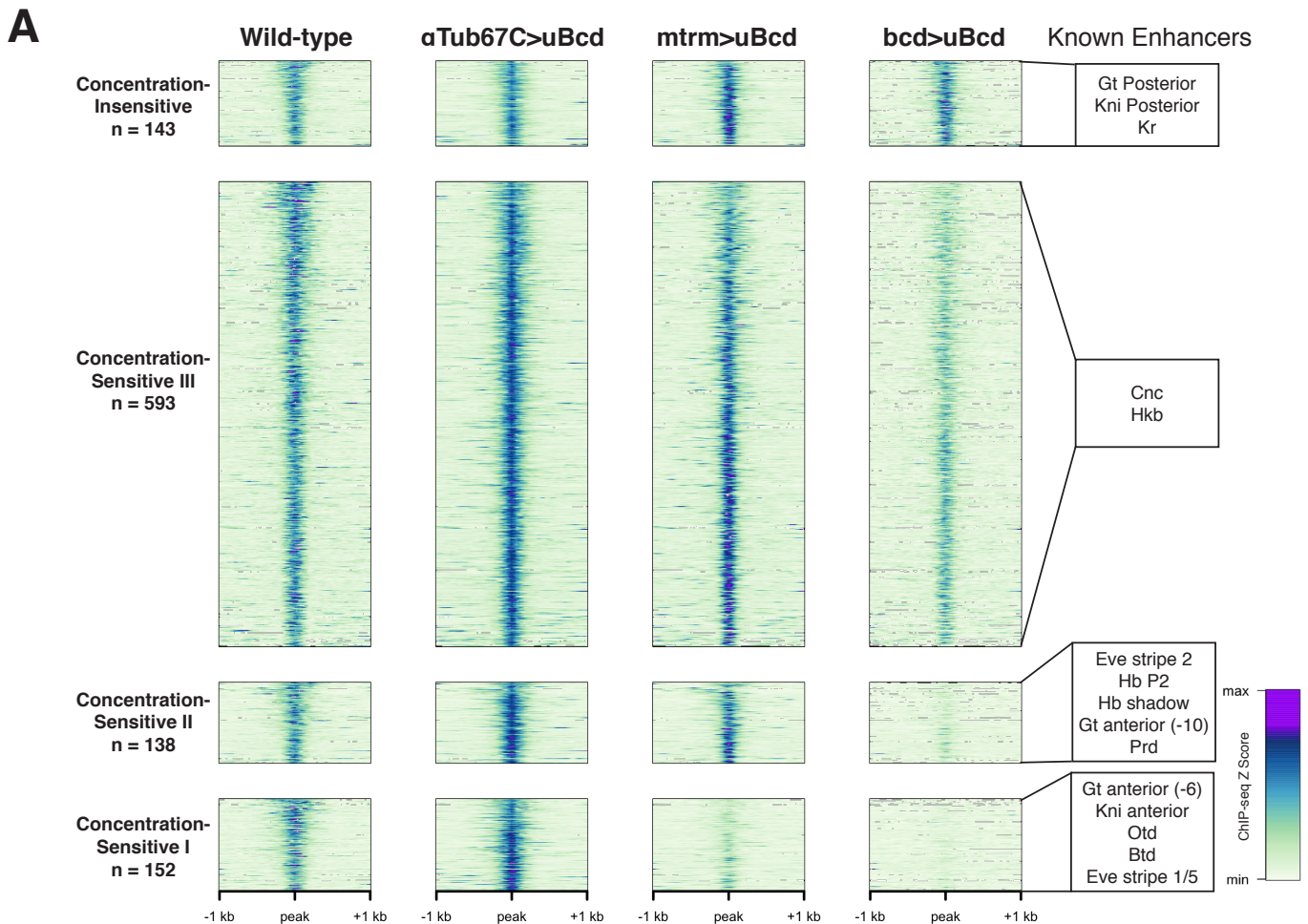
(B) Expression levels of uniform GFP-Bcd transgenic constructs relative to wild-type Bcd expression. Live embryos were imaged in during NC14, and dorsal profiles were plotted. Error bars are standard error of the mean. For wild-type, n = 23 embryos; *bcd*-uBcd n = 13; *mtrm*-uBcd n = 7; and  $\alpha$ Tub67C -uBcd n = 14. See also Figures S1 and S2 and Table S1.

(C) Immunostaining as (A), for each level of uniform Bcd. Anterior target gene expression is absent at the lowest level. At intermediate (*mtrm*) and high ( $\alpha$ Tub67C) levels of uBcd, anterior expression patterns are expanded and/or duplicated in the posterior, and posterior expression of Kni is absent.

(D) Larval cuticle preparations for the indicated genotypes. Embryos are oriented with anterior at the top. Head structures are indicated with open arrows and tail structures with filled arrows.  $\alpha$ Tub67C>uBcd embryos develop essentially no cuticle tissue, but form only what appear to be anteriorly-derived mouth structures. *mtrm*>uBcd results in a

duplication of the anterior-most abdominal denticles in the anterior and posterior of the embryo, with no clear terminal structures forming at either end. *bcd>uBcd* embryos have a normal posterior and all abdominal segments, but no thoracic or head structures. Images of individual embryos were rotated and cropped to exclude nearby embryos and air bubbles.





**Figure 2. Bcd-bound regions are classified into groups of increasing sensitivity to Bcd concentration.**

(A) ChIP-seq data in Bcd-bound peaks. Data is displayed as a heatmap of z-score normalized ChIP-seq reads, in a 2 kb region centered around each peak. Peaks in each class are arranged in order of decreasing z-scores in wild-type embryos. One peak (peak 549, see Table S6) was not classified, as it showed increasing binding at decreasing Bcd concentrations. Previously characterized enhancers overlapping with each class are indicated at right. Concentration-Insensitive: the posterior stripe enhancers for both *knirps* (Pankratz et al., 1992) and *giant* (Schroeder et al., 2004), and the *Kr CD1* enhancer (Hoch et al., 1991).

Concentration-Sensitive III: *cap'n'collar* (Schroeder et al., 2004), and *huckebein* (Häder et al., 2000) enhancers.

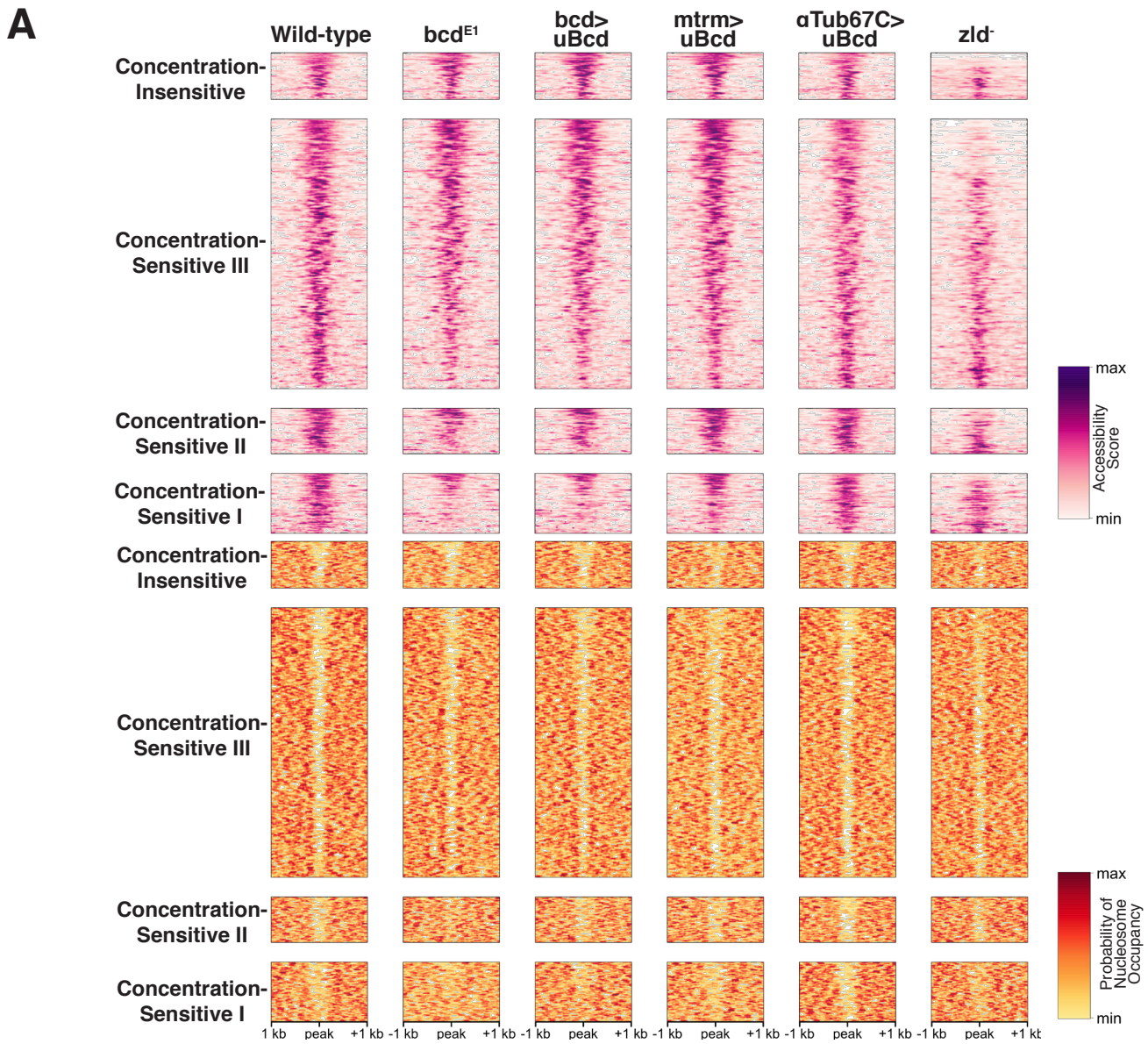
Concentration-Sensitive II: the *hunchback* P2 proximal (Struhl et al., 1989) and shadow enhancers (Perry et al., 2011), the *even-skipped* stripe 2 enhancer (Goto et al., 1989), an early *paired* enhancer (Ochoa-Espinosa et al., 2005), and an anterior enhancer for *giant* (Schroeder et al., 2004).

Concentration-Sensitive I: *buttonhead* (Wimmer et al., 1995), *orthodenticle* (Gao and Finkelstein, 1998), and anterior enhancers for both *knirps* and *giant* (Schroeder et al., 2004).

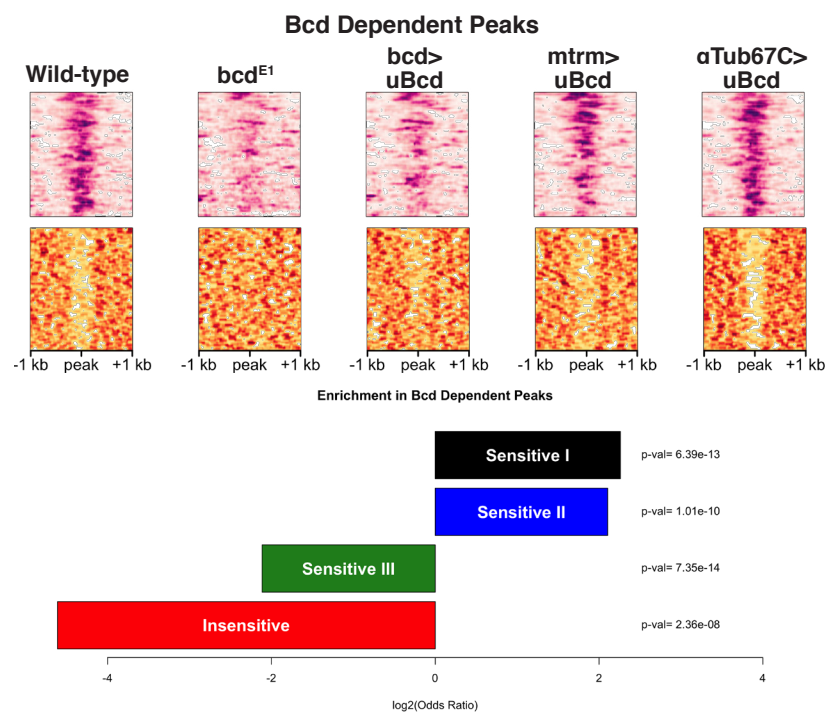
(B) Mean expression patterns of Vienna Tile-GAL4 enhancer reporters overlapping with Bcd peaks in each sensitivity class. Peaks and tiles with more than one overlap were excluded from the plot.

(C) Top DNA motifs discovered by RSAT peak-motifs. The e-value for is a p-value computed from a binomial distribution for a given motif in the dataset, corrected for multiple testing. See Figure S2B for *de novo* motif discovery in each sensitivity class.

(D) Heatmap displaying the enrichment of a given motif in each sensitivity class, relative to the peak list as a whole. P-values were generated from permutation tests (n = 10,000 tests).



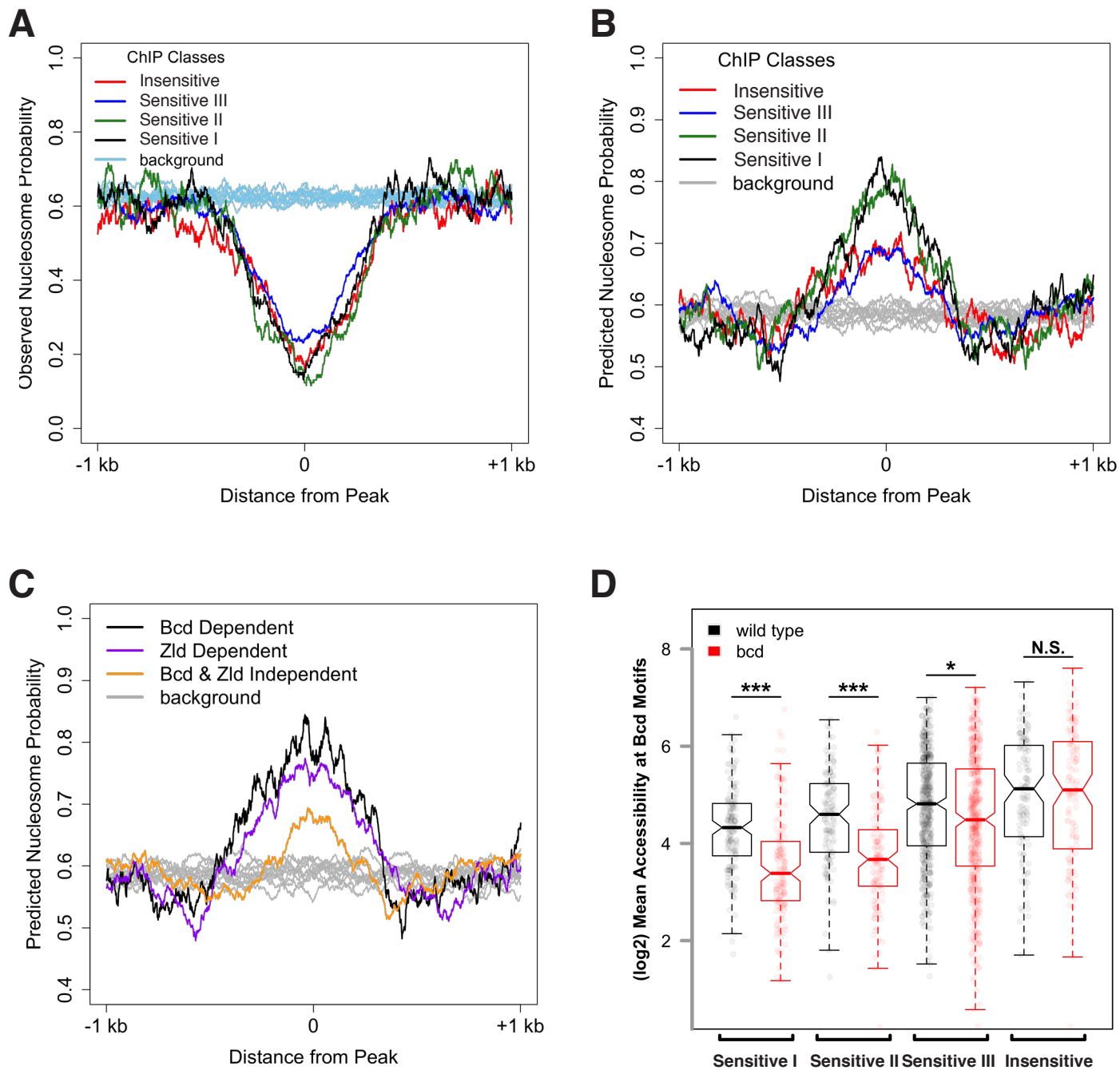
**B**



**Figure 3. Bcd drives chromatin accessibility primarily at concentration-sensitive targets.**

(A) Heatmaps showing chromatin accessibility (top) and probability of nucleosome occupancy (bottom) around Bcd-bound peaks from ATAC-seq experiments. Peak regions are arranged by decreasing accessibility in wild-type embryos. *bcd<sup>E1</sup>* mutant embryos show a loss of accessibility and increased nucleosome occupancy most strongly at the Concentration-Sensitive I and II peaks. *zld* embryos show reduced accessibility across all sensitivity classes.

(B) Subset of 132 Bcd-bound peaks selected from (A) that become inaccessible in the absence of Bcd. Accessibility at these peaks increases with increasing concentrations of uniform Bcd. Odds ratios and p-values calculated from Fisher's exact test show significant overrepresentation of the Concentration-Sensitive I and II classes in the Bcd-dependent peaks.



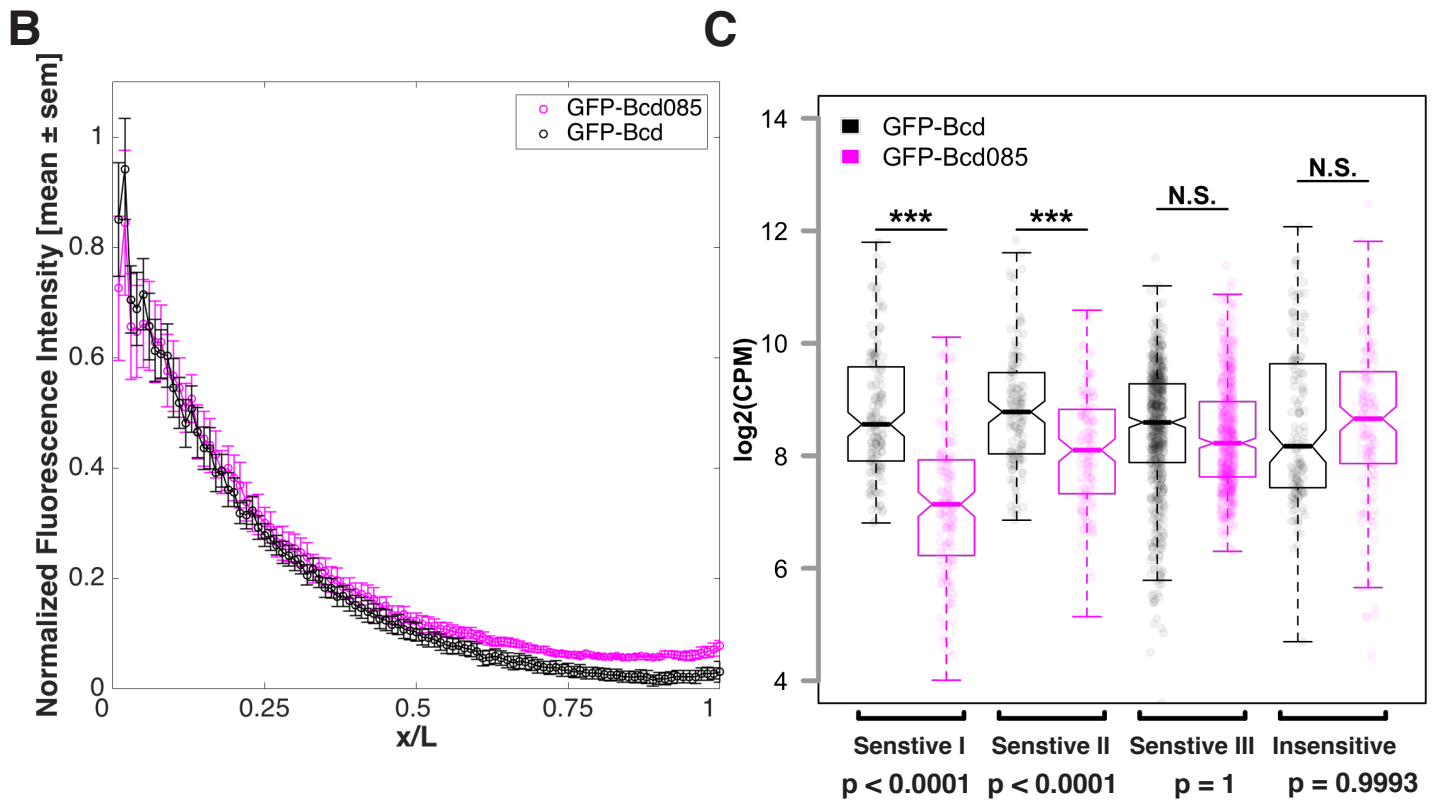
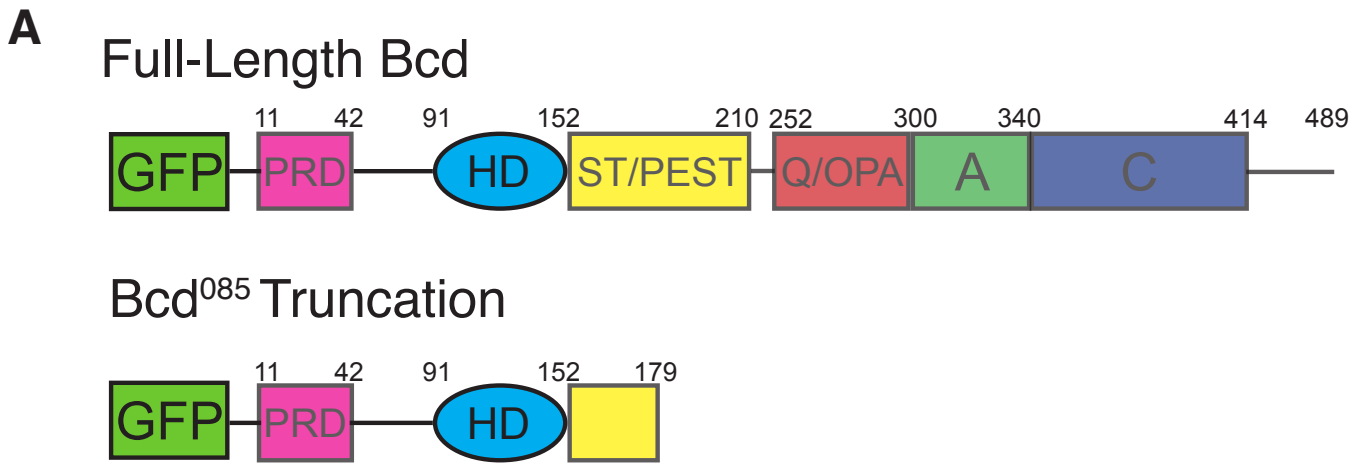
**Figure 4. Bcd sensitivity classes differ in both predicted and observed nucleosome occupancy.**

(A) Metaprofiles of nucleosome occupancy in each sensitivity class in wild-type embryos. Background represents random selection of regions outside of Bcd peaks shows a genome-wide average nucleosome probability of  $\sim 0.6$ . Bcd-bound peak regions show reduced nucleosome occupancy compared to unbound regions.

(B) Predicted nucleosome occupancy using NuPoP show higher modeled probability of nucleosome occupancy in Bcd-bound peaks relative to background regions, with higher probability of occupancy at the Concentration-Sensitive I and II classes.

(C) Predicted nucleosome occupancy in peaks dependent on Bcd vs. Zld ( $n_{\text{Bcd}}=132$  peaks,  $n_{\text{Zld}}=402$  peaks, with  $n=61$  peaks dependent on both Bcd and Zld) for accessibility show higher predicted occupancy than peaks independent of both Bcd and Zld ( $n=554$ ).

(D) Mean wild type (black) or *bcd*<sup>-</sup> (red) ATAC accessibility scores for Bcd motifs were calculated for each peak and plotted by sensitivity group. Boxplots depict the distribution of accessibility scores for each group in each genotype, and individual data points are shown as points. P-values were calculated by one-sided permutation test and indicate the likelihood in a randomly selected population of observing a difference between means greater than the observed values ( $p < 1e-6$  for Concentration-Sensitive I and II groups,  $p = 0.001207$  for Concentration-Sensitive III, and  $p = 0.988167$  for Concentration-Insensitive).



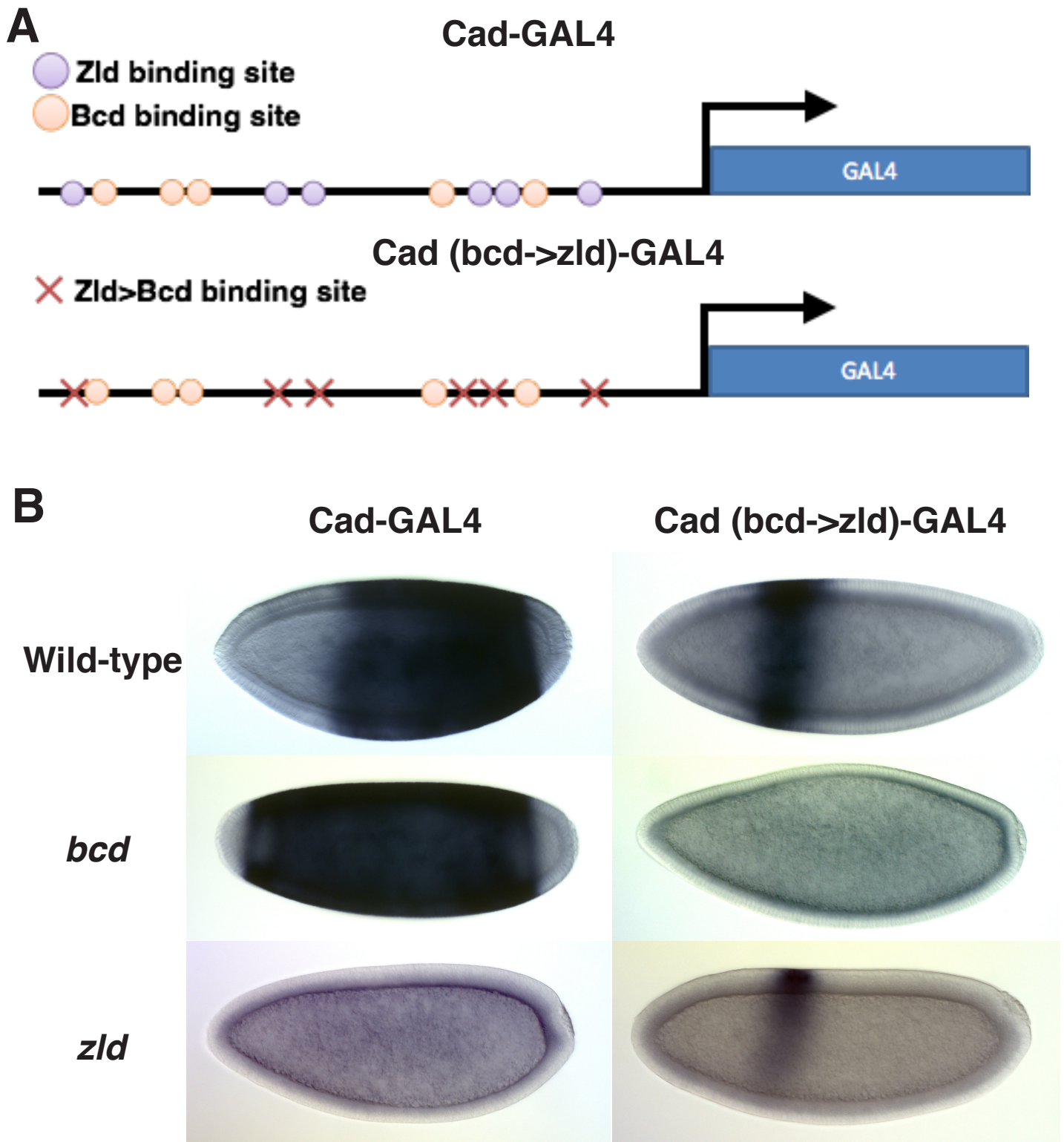


**Figure 5. Bicoid requires C-terminal protein domains to bind to concentration-sensitive targets.**

(A) GFP-Bcd<sup>085</sup> construct is truncated within the S/T domain downstream of the home domain. Wild-type protein domains modified from (Janody et al., 2001) and (Crauk and Dostatni, 2005). The N-terminus of the protein includes a PRD repeat, followed by the DNA-binding homeodomain (HD) (Berleth et al., 1988). The serine/threonine-rich (S/T) domain is the target of MAPK phosphorylation by the terminal patterning Torso pathway (Janody et al., 2000) and contains a PEST sequence implicated in targeting the protein for degradation (Rechsteiner and Rogers, 1996). The C-terminus contains three domains implicated in transcriptional activation. The glutamine-rich (Q)/OPA and alanine-rich (A) domains are required for interactions with TAFII110 and TAFII60, respectively (Sauer et al., 1995). The acidic (C) domain has been demonstrated to play a role in transcriptional activation in yeast, but is not required for Bicoid activity in the embryo (Driever et al., 1989a).

(B) GFP-Bcd<sup>085</sup> forms a protein gradient comparable to wild-type GFP-Bcd. GFP fluorescence intensity was extracted from dorsal profiles of live embryos. Error bars are standard error of the mean: GFP-Bcd embryos, n = 8; and GFP-Bcd<sup>085</sup> embryos, n = 8.

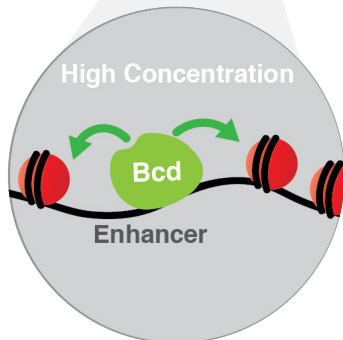
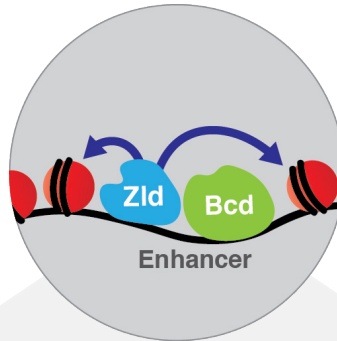
(C) Boxplots displaying log transformed CPM normalized ChIP-seq data from GFP-Bcd;;bcd<sup>E1</sup> (wild-type) and GFP-Bcd<sup>085</sup>;;bcd<sup>E1</sup> (Bcd<sup>085</sup>) embryos show significant reduction binding of Bcd<sup>085</sup> in Concentration-Sensitive I and II peaks. P-values were calculated from permutation tests (n = 10,000). See also Figure 2 Figure Supplement 2.



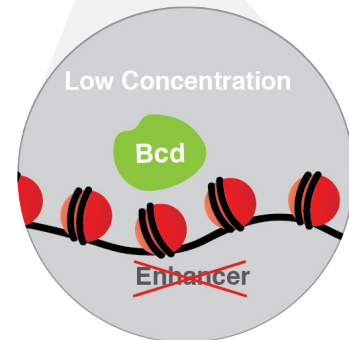
**Figure 6. Replacing Zld sites with Bcd sites shifts gene expression to the anterior.**

- (A) Schematic of the Vienna Tile enhancer reporter for *caudal*, containing 5 Zld and 6 Bcd binding sites. The mutated reporter contains 11 Bcd binding sites and no Zld sites.
- (B) Expression of the wild-type and mutated reporter in wild-type, *bcd*<sup>-</sup> or *zld*<sup>-</sup> embryos.

# Concentration Insensitive

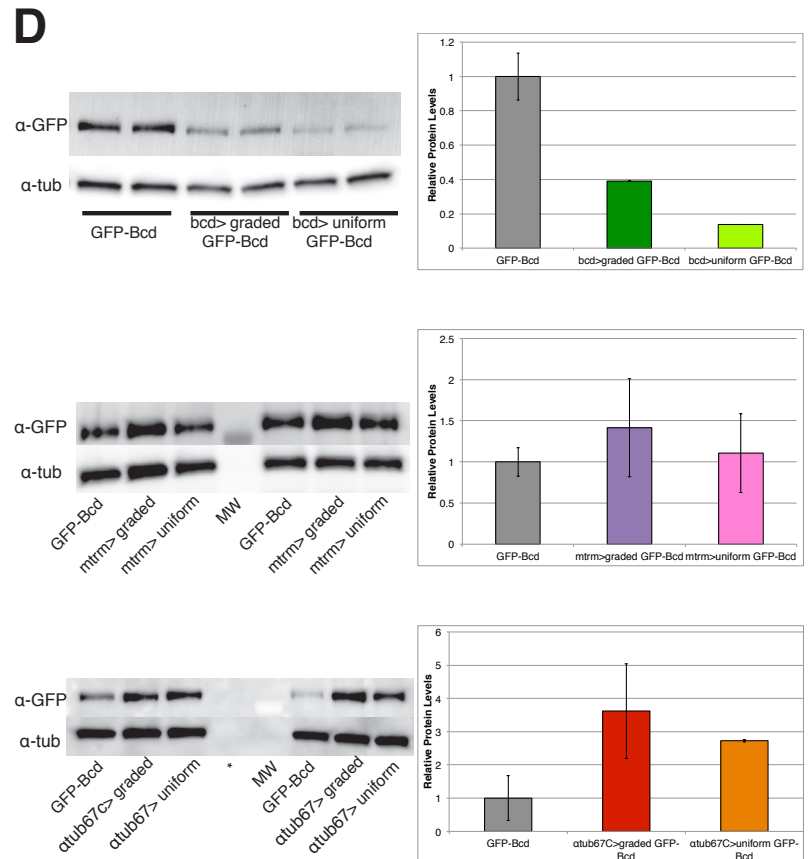
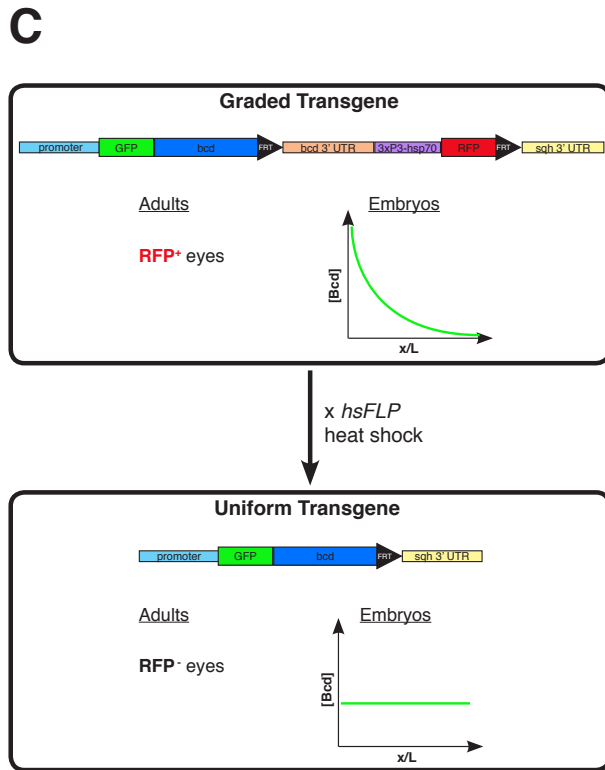
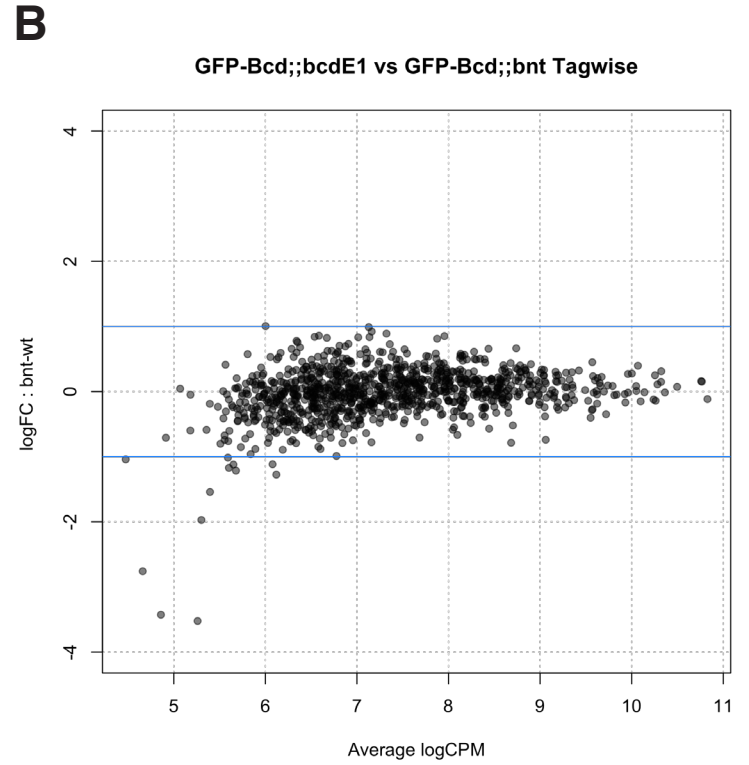
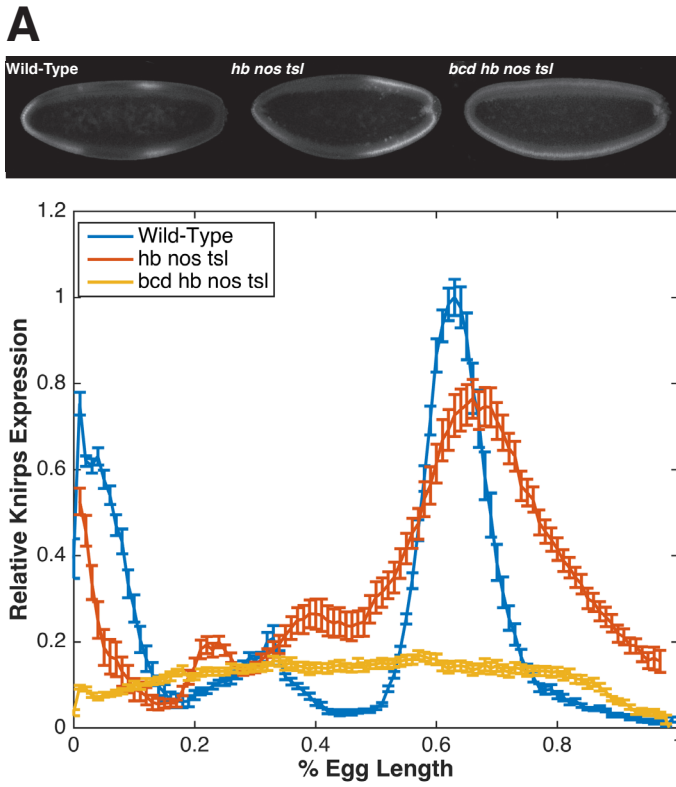


# Concentration Sensitive



### **Figure 7. Model for Bicoid function along the AP axis.**

Bcd drives accessibility of concentration-sensitive, Bcd-dependent enhancers at high concentrations in anterior nuclei, and these sites are closed in posterior nuclei. concentration-insensitive targets remain accessible in both anterior and posterior nuclei, likely through inputs from other factors such as Zld and more open local chromatin structure with a lower nucleosome preference.



## Figure 1 Figure Supplement 1.

A) Quantification of Knirps intensity in wild-type, triple mutant (*hb nos tsl*) and quadruple mutant (*bcd hb nos tsl*). Bicoid activates patterned expression of Knirps. In embryos in which Bicoid is the only source of maternal patterning information (*hb nos tsl*), a broad domain of Kni is expressed in the posterior of the embryo. In quadruple mutant embryos, a low level of uniform Knirps is expressed ubiquitously, suggesting that Bcd is required for activating expression of *knirps* above a background level. Heat-fixed embryos from wild-type (Oregon-R) mothers, *hunchback nanos torso-like* germline clones and *bicoid hunchback nanos torso-like* germline clones were pooled and immunostained in a single tube with a rat anti-Knirps primary antibody and Alexa-647 rat antibody. Embryos were mounted on a single slide and imaged by confocal microscopy. Representative embryos for each genotype are shown. Fluorescence intensity of Knirps was extracted from dorsal profiles of midsagittal sections of embryos and plotted using MATLAB. Data are fluorescence intensity minus background, and error bars are standard error of the mean for  $n = 5$  wild-type,  $n = 8$  *hb nos tsl*, and  $n = 6$  *bcd hb nos tsl* embryos.

B) Smear plot generated in EdgeR (Robinson et al., 2010) showing the log transformed fold-change in Bcd binding between mutant and wild-type embryos for each Bcd peak, vs. the average log transformed sequencing read counts per million (CPM). Bcd binding shows no significant changes between wild-type and *nos tsl* mutant embryos.

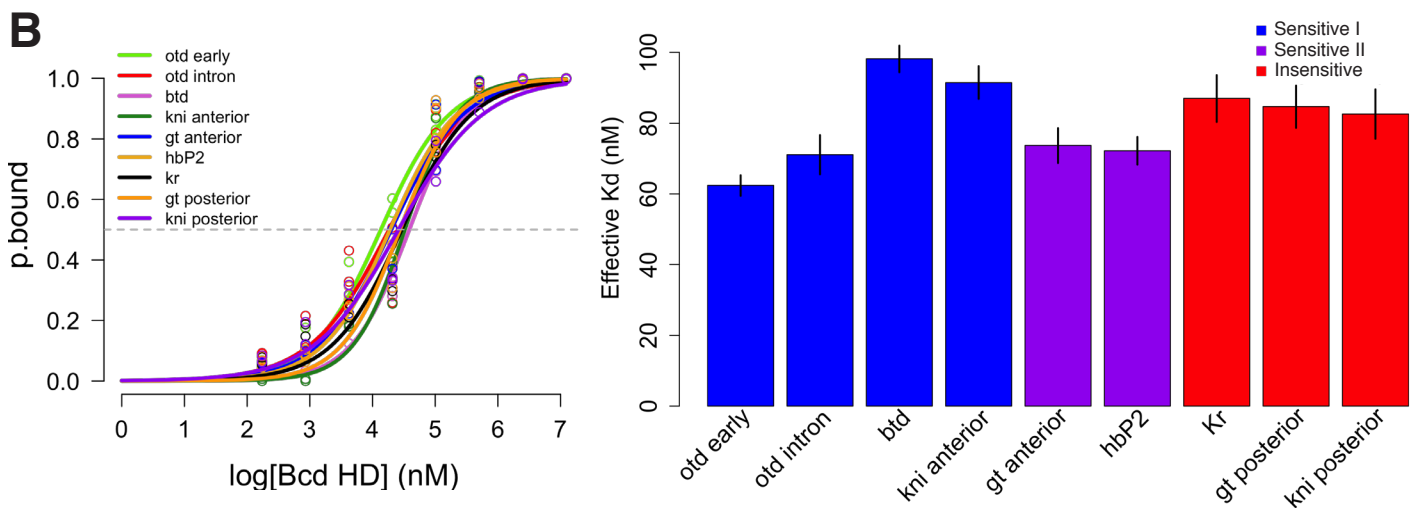
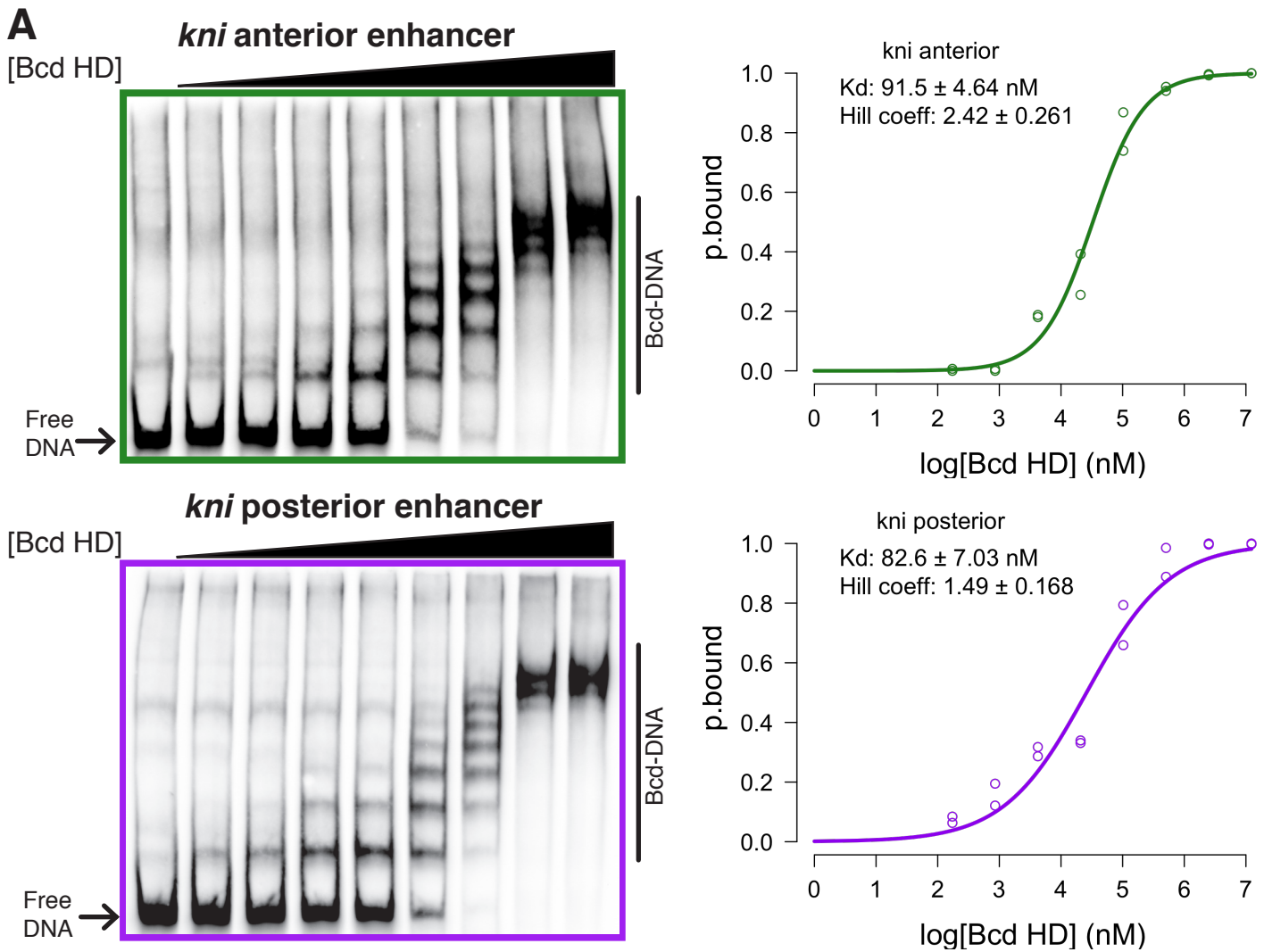
Significance was determined using EdgeR to perform a pairwise exact test with a cutoff of  $FDR \leq 0.05$ , comparing binding between *eGFP-Bcd;;bcd<sup>E1</sup>* and *eGFP-Bcd;;bcd<sup>E1</sup>hb<sup>FB</sup>nos<sup>L7</sup>tsl<sup>A</sup>* in the 1,027 Bcd peaks.

C) Schematic of the uniform Bcd transgene. The uniform Bcd transgene contains an N-terminal GFP-tagged Bcd driven by the various maternal promoters discussed in the text. Downstream of the *bcd* coding sequence is a cassette containing the endogenous *bcd* 3'UTR and a 3xP3-hsp70 promoter driving promoter of RFP. This cassette is flanked by FRT sites. The *sqh* 3'UTR lies downstream of the FRT cassette. Flies expressing this version of the transgene can be identified by RFP expression in their eyes, and females produce embryos in which Bcd is distributed in a gradient. Males from this transgenic stock are crossed to females expressing a heat shock inducible flippase (*hsFLP*), and heat shocking the F1 larvae results in recombination and excision of the cassette at the FRT sites, bringing the *sqh* 3'UTR directly downstream of the *bcd* coding sequence. This initially results in mosaic F1 flies with a mosaic graded/uniform Bcd germline. The F1 are further outcrossed to *bcd*<sup>E1</sup> mutants and F2 individuals producing embryos with uniform Bcd distributions can be identified by the lack of RFP expression in the eyes.

D) Expression levels of uniform Bcd constructs measured by western blots.

Western blots for GFP-Bcd were performed on embryos at NC14. Representative gels and quantifications are shown for the *bcd* promoter-driven transgene (A), *mtrm* promoter-driven transgene (B) and  $\alpha$ -tub67C promoter-driven transgene (C). In the barplots, band intensities are reported relative to wild-type (GFP-Bcd). All lanes are normalized to an  $\alpha$ -tubulin loading control. Error bars are standard deviation between two biological replicates for each sample. MW = molecular weight marker, \* = skipped lane.



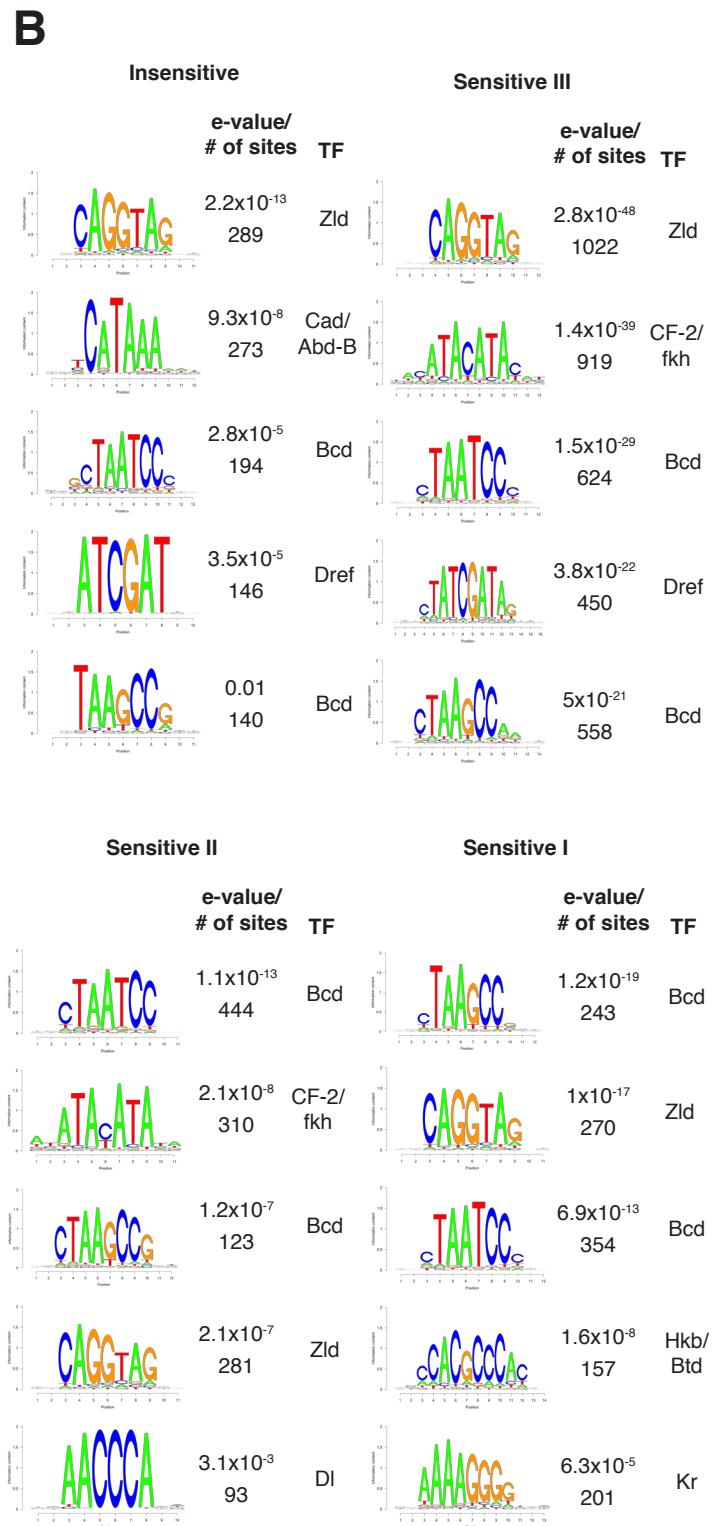
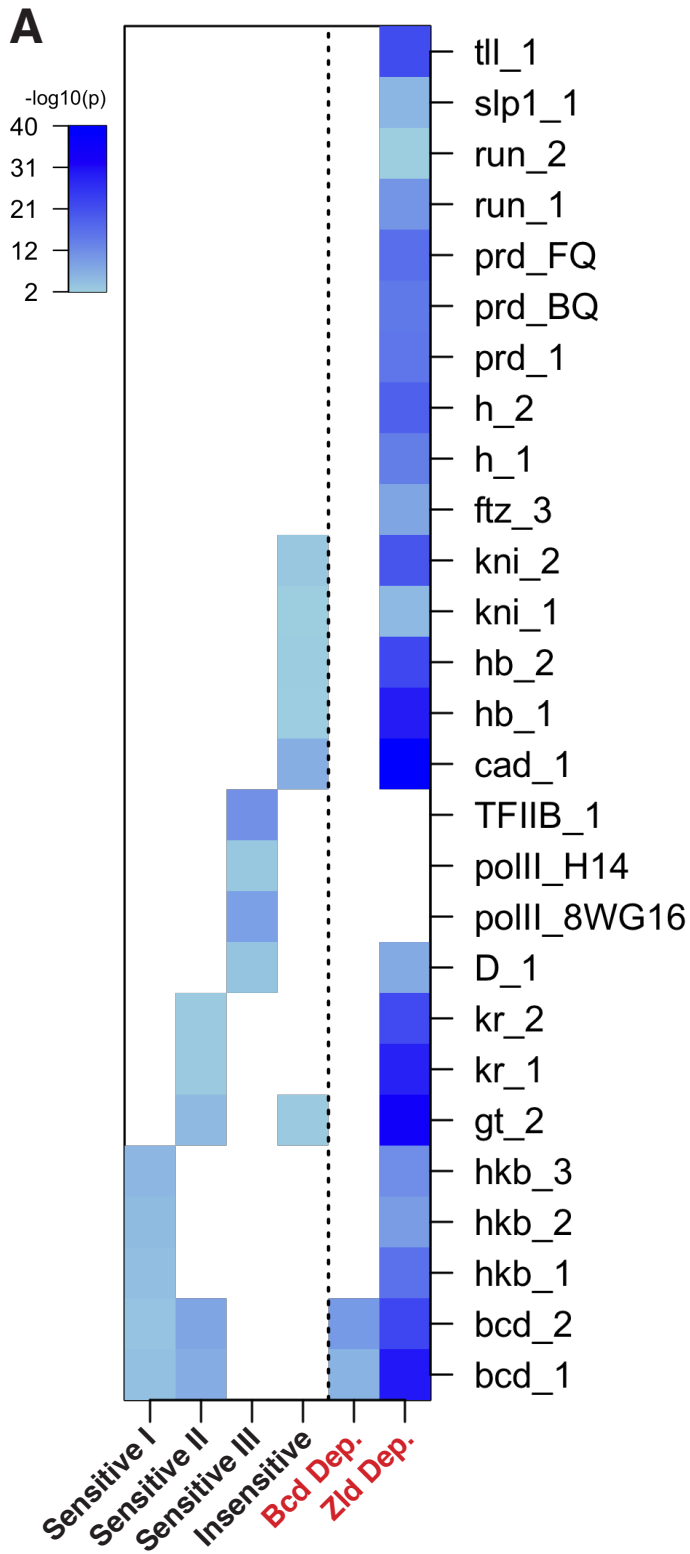


**Figure 2 Figure Supplement 1. *In vitro* binding affinity of target enhancers for Bcd protein is insufficient to explain *in vivo* binding behavior.**

(A) Representative gels from EMSAs with *kni* anterior or posterior enhancer sequence used as DNA probe. Binding curves display the log transformed Bcd concentration is plotted vs. ratio of bound to shifted probe (p.bound).

(B) Binding curves for nine EMSA probes show largely overlapping profiles of *in vitro* affinity for Bcd.

(C) Effective  $K_d$  measurements for nine EMSA probes do not correspond to *in vivo* behavior of the same DNA sequences. *In vivo* sensitivity classifications determined by ChIP-seq are indicated by color of bars. Error bars are standard error from 2-3 technical replicates per DNA probe.



**Figure 2 Figure Supplement 2. Enrichment for binding and motifs of transcription factors in Bcd sensitivity classes.**

(A) Heatmap depicting enrichment of Berkeley Drosophila Transcription Network Project (BDTNP) ChIP-chip peaks for AP factors in Bcd-bound sensitivity classes. ChIP data (MacArthur et al., 2009) was downloaded from [bdtnp.lbl.gov](http://bdtnp.lbl.gov), and overlap between ChIP peaks for the indicated factors/antibodies and either the Bcd sensitivity classes or ATAC-seq accessibility dependence groups were calculated. One-sided Fisher's exact tests were performed to test for enrichment of a BDTNP ChIP peak set within given Bcd peak class. P-values are plotted as  $-\log_{10}$  values, where white indicates non-significant values.

(B) *De novo* motif discovery performed with RSAT as in Figure 2B, for each of the Bcd sensitivity classes individually. The top five enriched motifs are displayed for each sensitivity class.