

1 **Neuroblast-specific chromatin landscapes allow integration of spatial and temporal cues**  
2 **to generate neuronal diversity in Drosophila**

3  
4 Sonia Q Sen<sup>1</sup>, Sachin Chanchani<sup>1</sup>, Tony D Southall<sup>2</sup>, and Chris Q Doe<sup>1\*</sup>

5  
6 1 Institute of Neuroscience, Institute of Molecular Biology, Howard Hughes Medical Institute,  
7 University of Oregon, Eugene, OR 97403

8 2 Department of Life Sciences, Sir Ernst Chain Building, Imperial College London, London,  
9 SW7 2AZ, United Kingdom

10  
11 \*Author for correspondence at [cdoe@uoregon.edu](mailto:cdoe@uoregon.edu)

12  
13  
14  
15 **Abstract**

16 During early neurogenesis in flies and mice, spatial and temporal cues interact to specify  
17 neuronal diversity, yet in no organism is it known how spatial and temporal cues are integrated.  
18 We used Targeted DamID (TaDa) to identify the genomic binding sites of the temporal  
19 transcription factor Hunchback in two adjacent Drosophila neuroblasts (NB5-6 and NB7-4).  
20 Hunchback targets were different in each neuroblast. Profiling chromatin accessibility showed  
21 that each neuroblast had a distinct chromatin landscape: Hunchback-bound loci in NB5-6 were  
22 in open chromatin, but the same loci in NB7-4 were in closed chromatin. Moreover, binding of  
23 the spatial factor Gsb/Pax3, essential for NB5-6 specification, was correlated with open  
24 chromatin and Hunchback-enriched loci in NB5-6, but not NB7-4. We propose early-acting  
25 spatial factors establish a unique chromatin landscape in each neuroblast, thereby restricting  
26 temporal factor binding to different loci in each neuroblast, resulting in different neurons in each  
27 neuroblast lineage.

28  
29 **Impact statement**

30 Integration of spatial and temporal identity during Drosophila neurogenesis is due to spatial factors  
31 generating neuroblast-specific chromatin thereby biasing subsequent temporal transcription factor  
32 binding and producing neuroblast-specific neurons.

33

## 34 Introduction

35  
36 The generation of neuronal diversity in mammals and *Drosophila* is a multi-step process. The  
37 initial step is the production of the neuroectoderm (ventral in *Drosophila*, dorsal in mammals) that  
38 gives rise to neural progenitors. In both systems, the neuroectoderm and neural progenitor  
39 population acquire regional differences due to the action of Hox genes and spatial patterning  
40 genes (Jessell, 2000). Although spatial patterning generates diversity within the neural progenitor  
41 population, it is insufficient to account for the neuronal diversity in the mature nervous system.  
42 Expanding neural diversity requires a second step called temporal patterning, where individual  
43 neural progenitors produce a sequence of distinct neurons and glia (Doe, 2017). In both  
44 *Drosophila* and mammals, this process appears to be regulated, in part, by temporal  
45 transcription factors (TTFs) that are sequentially expressed within individual neural progenitors  
46 (Kohwi and Doe, 2013). Although a great deal is known about how spatial factors generate  
47 regional diversity, and much has recently been learned about temporal patterning mechanisms,  
48 virtually nothing is known about how spatial factors and TTFs are integrated to specify distinct  
49 neuronal identities in spatially distinct progenitor populations.

50 *Drosophila* is an excellent model system to investigate how spatial and temporal factors are  
51 integrated during neurogenesis, due to a deep understanding of neural progenitor (neuroblast)  
52 lineages, and the molecular mechanisms involved in both spatial and temporal patterning during  
53 neurogenesis. The *Drosophila* neuroectoderm produces a bilateral array of 30 neuroblasts in  
54 each segment, named according to their row and columnar position within the two dimensional  
55 neuroblast array (Figure 1A, left). Each neuroblast has a unique identity based on its distinct  
56 molecular profile and each neuroblast produces a unique and stereotyped family of neurons.

57 Spatial patterning factors that specify neuroblast identity have been characterized, and all of  
58 them are transcription factors or signalling pathways with transcription factor effectors.  
59 Henceforth we refer to these spatial factors as “spatial transcription factors” or STFs, paralleling  
60 the naming of temporal transcription factors as TTFs. The Gooseberry (Gsb) Pax-3 family  
61 transcription factor is expressed in row 5 neuroblasts; loss of Gsb transforms row 5 neuroblasts  
62 into row 3/4 identity, and misexpression of Gsb transforms row 3/4 neuroblasts into row 5  
63 identity. Importantly, transient misexpression of Gsb in the neuroectoderm, prior to neuroblast  
64 formation, is sufficient to generate ectopic row 5 neuroblasts, suggesting that neuroblast identity  
65 is determined in the neuroectoderm and maintained during the subsequent neuroblast lineage  
66 (Bhat, 1996; Skeath et al., 1995). Thus, Gsb is one of the best characterized STFs. Similarly, the  
67 secreted Wingless (Wg) protein is produced by row 5 neuroectoderm, where it is required to  
68 specify the adjacent row 4 and 6 neuroblast identity that is maintained in the row 4 and 6  
69 neuroblasts (Chu-LaGraff and Doe, 1993). Precise inactivation of a temperature-sensitive Wg  
70 protein showed that loss of Wg activity in the neuroectoderm resulted in loss of neuroblast  
71 identity, whereas inactivation of Wg after neuroblast formation had no effect, showing that  
72 transient Wg generates row 4 and 6 neuroblast identity (Chu-LaGraff and Doe, 1993). In addition,  
73 Hedgehog (Hh) expression in row 6/7 neuroectoderm is required to specify neuroblast identity in  
74 adjacent rows 1/2 (McDonald and Doe, 1997). Finally, Engrailed expression in the neuroectoderm  
75 is required for the proper development of row 6/7 neuroblasts, and transient Engrailed  
76 misexpression generates ectopic row 7 neuroblast identity (Deshpande et al., 2001). Taken

77 together, these spatial patterning experiments show that neuroblast spatial identity is specified in  
78 the neuroectoderm by the transient action of STF s expressed in different neuroblast rows.

79 Spatial patterning does not only generate distinct rows of neuroblasts, but also distinct  
80 neuroblast columns. During the first stages of neuroblast formation there are three distinct  
81 columns of neuroblasts, each specified by a conserved homeodomain protein. Vnd is expressed  
82 in a medial column of neuroectoderm, Ind is expressed in an intermediate column, and Msh  
83 (Flybase: Drop) is expressed in the lateral column (Figure 1A, left) (Isshiki et al., 1997; McDonald  
84 et al., 1998; Weiss et al., 1998). Loss of function and misexpression studies show that each is  
85 necessary and partially sufficient for specifying columnar neuroblast identity (Isshiki et al., 1997;  
86 McDonald et al., 1998; Weiss et al., 1998). It is likely that these columnar factors function in the  
87 neuroectoderm, like spatial row factors, because they do not persist throughout neuroblast  
88 lineages. All three of these STF s have conserved mammalian orthologs with similar medial-lateral  
89 expression in the neuroectoderm (Weiss et al., 1998). Overall, the combination of row and  
90 columnar STF s are likely to generate the observed 30 distinct neuroblast identities. Hox factors  
91 provide an additional spatial cue that distinguishes segmental differences in neuroblast identity  
92 (Prokop and Technau, 1994).

93 Whereas spatial patterning generates 30 different neuroblast identities, temporal patterning is  
94 required to generate different progeny within each neuroblast lineage. Most neuroblasts  
95 sequentially express a series of four TTF s as they divide to generate ganglion mother cell (GMC)  
96 progeny, and the specific TTF inherited by each GMC determines its identity (Doe, 2017; Kohwi  
97 and Doe, 2013; Li et al., 2013). Embryonic ventral nerve cord (VNC) neuroblasts undergo a TTF  
98 cascade that progresses from Hunchback (Hb; Ikaros zinc finger family) to Krüppel (zinc finger  
99 family) to the redundant Nubbin/Pdm2 (Pdm) to Castor (Cas; Casz1 zinc finger family) (Figure 1A,  
100 middle). Other neuroblasts in the larval VNC, brain, and optic lobes undergo a similar TTF  
101 cascade to increase neuronal diversity, although the identity of the TTF s differ in each region  
102 (Doe, 2017; Li et al., 2013). The Hb-Kr-Pdm-Cas TTF cascade has been particularly well-  
103 characterized, with each factor being necessary and sufficient to specify the neuronal identity  
104 produced during its window of expression (Grosskortenhau s et al., 2006; Isshiki et al., 2001;  
105 Kanai et al., 2005; Kohwi et al., 2013; Novotny et al., 2002; Tran and Doe, 2008). Importantly,  
106 each TTF specifies a different type of neuron in each neuroblast lineage, showing that spatial  
107 identity provides a different context for Hb function in each neuroblast (Figure 1A, right).  
108 Understanding this “context” at a mechanistic level is the goal of our experiments below.

109 The role of TTF s is best exemplified by Hb, the first TTF in the cascade. Loss of Hb results in  
110 absence of the first-born neuron identities in all neuroblast lineages assayed to date (1-1, 3-1, 3-  
111 5, 7-1, 7-3). Conversely, driving prolonged Hb expression in neuroblasts results in ectopic first-  
112 born neurons in all lineages tested (Isshiki et al., 2001; Kanai et al., 2005; Kohwi et al., 2013;  
113 Novotny et al., 2002). For example, prolonged expression of Hb in NB7-1 produces ectopic U1  
114 motor neurons, whereas prolonged expression of Hb in NB7-3 produces ectopic EW1  
115 serotonergic interneurons. Note that these misexpression experiments further confirm the  
116 neuroblast-specific effect of Hb, showing that the spatial identity of the neuroblast determines  
117 the effect of Hb. Importantly, Hb can induce early-born neuronal identity throughout a  
118 “competence window” of ~5 neuroblast divisions (from embryonic stage 9-12). The length of the  
119 competence window is defined by expression of Distal antenna (Dan), a nuclear Pipsqueak

120 domain protein present in all neuroblast nuclei until stage 12 (about five divisions for most  
121 neuroblasts); Dan is downregulated in all neuroblasts at the end of stage 12, and this closes the  
122 Hb competence window (Kohwi et al., 2013). Hb can induce first-born neuronal identity at any  
123 point during this competence window, showing that Hb binding sites are accessible throughout  
124 the competence window; this is important to consider for the experiments described here, where  
125 we have restricted our Hb binding and chromatin accessibility profiling experiments to the stage  
126 9-12 competence window in individual neuroblast lineages (see below).

127 It is clear that spatial and temporal cues are integrated to generate lineage-specific neuronal  
128 diversity, both in *Drosophila* embryonic neuroblasts and optic lobe neuroblasts (Erclik et al.,  
129 2017), and likely in mammalian progenitor lineages. Yet in no case, mammals or *Drosophila*, is it  
130 known how spatial and TTFs are integrated. Here we hypothesise two mechanisms by which this  
131 integration could occur. (1) Independent specification (Figure 1B). In this scenario, spatial and  
132 temporal transcription factors bind their genomic targets independently, and the combinatorial  
133 actions of these factors and their downstream gene regulatory networks results in unique gene  
134 expression and therefore unique neural identities. (2) Sequential specification (Figure 1C). In this  
135 scenario, early expression of STFs in the neuroectoderm (where they are known to act) biases  
136 the subsequent DNA-binding profile of the later expressed TTFs. This could happen via STFs  
137 generating different chromatin landscapes in each neuroblast, or via STFs promoting the  
138 persistent expression of TTF cofactors that result in neuroblast-specific TTF DNA-binding. While  
139 both scenarios would result in the specification of distinct neural identities in spatially distinct  
140 NBs, in the independent specification model, TTF binding will be identical in all neuroblasts  
141 whereas in the sequential specification model, TTF binding will occur at different loci in each  
142 neuroblast.

143 To discriminate between these models, we sought to determine Hb genomic targets in NB5-  
144 6 versus NB7-4. If independent specification is used, we expect to find similar Hb occupancy in  
145 each neuroblast (Figure 1B), whereas if sequential specification is used, we expect to find  
146 different Hb genomic binding in each neuroblast (Figure 1C). Our goal was to identify Hb  
147 occupancy within the early NB5-6 and NB7-4 lineages during the Hb competence window, when  
148 Hb retains the ability to generate ectopic early-born neuronal identities, and thus presumably can  
149 still bind its normal genomic targets. To identify Hb occupancy in these two neuroblast lineages,  
150 we adapted the previously described TaDa method (Marshall et al., 2016; Southall et al., 2013).  
151 TaDa relies on an attenuated expression of the DNA adenosine methyltransferase (Dam) enzyme  
152 (Figure 1D), which binds genomic DNA and methylates adenosine at GATC sites. This covalent  
153 DNA mark can be used to determine Dam binding sites, due to the very low level of endogenous  
154 DNA methylation in *Drosophila*. Expression of Dam alone can be used to detect open chromatin  
155 (Aughey et al., 2018) (Figure 1E) or Dam can be fused to a transcription factor such as Hb, which  
156 provides a read-out of Hb genomic occupancy (Figure 1F).

157 Here we characterize two Gal4 lines that are specific for NB5-6 and NB7-4 lineages in the  
158 embryo. We use these lines to obtain NB-specific expression of Dam:Hb (to identify Hb genomic  
159 occupancy) and Dam alone (to detect open chromatin). We demonstrate that Hb has differential  
160 targets in NB5-6 and NB7-4 lineages, which correspond to differentially open chromatin in each  
161 lineage. Importantly, our observation that Hb-bound loci specific to NB5-6 have open chromatin,  
162 but the same loci in NB7-4 have closed chromatin, shows that Hb is not sufficient to create open



163 chromatin. Rather, Hb binding in each neuroblast is likely restricted to a subset of neuroblast-  
164 specific open chromatin domains. In support of this model, the Gsb STF, required to specify  
165 NB5-6 but not NB7-4, shows enriched occupancy at open chromatin and Hb enriched loci in  
166 NB5-6, but not in NB7-4, consistent with a role for Gsb in generating neuroblast-specific open  
167 chromatin organization. Our findings support a sequential specification model in which STFs  
168 create neuroblast-specific chromatin organization, leading to neuroblast-specific Hb DNA-  
169 binding.

170

## 171 Results

172

### 173 Characterization of Gal4 lines specifically expressed in NB5-6 and NB7-4

174 Here we characterize two Gal4 lines that label either the NB5-6 or the NB7-4 lineages, which  
175 is a prerequisite for profiling neuroblast-specific Hb binding sites. NB5-6 forms in the Gsb  
176 domain, whereas NB7-4 forms in the Engrailed domain (Figure 2A). To label NB5-6 and its  
177 lineage we used *ladybird early (lbe)-Gal4*, which is reported to specifically label NB5-6 and its  
178 progeny (Baumgardt et al., 2009; Urbach and Technau, 2003). We confirmed that *lbe-Gal4*  
179 expression was highly specific to the NB5-6 and its lineage from stage 10 through stage 12, the  
180 time frame of our experiments (Figure 2B-D'; Figure 2 – Supplement 1A), although by stage 17 it  
181 has expression in the non-neuronal salivary gland (Figure 2 –Supplement 1A). Henceforth we call  
182 this line “NB5-6-Gal4.” To label NB7-4 and its lineage, we used the previously described  
183 *R19B03<sup>AD</sup> R18F07<sup>DBD</sup>* split-Gal4 line (Lacin and Truman, 2016). We confirmed that this line labels  
184 NB7-4 and its lineage from stage 10 until the end of stage 17 (Figure 2E-G'; Figure 2 –  
185 Supplement 1B); the only off-target expression is in the adjacent NB5-6 lineage in 6% of  
186 hemisegments (n=1176). Henceforth we call this line “NB7-4-Gal4.” Both *NB5-6-Gal4* and *NB7-4-Gal4*  
187 lines are first expressed after Hb expression in the NB, but during the ‘Hb competence  
188 window’ defined by the presence of Distal antenna (Dan) nuclear protein in stage 9-12  
189 neuroblasts (Figure 2C' and F') (Kohwi et al., 2013). Importantly, ectopic Hb can induce early-  
190 born neuronal identity throughout the Hb competence window, and thus the relevant Hb DNA-  
191 binding sites are still accessible. We conclude that *NB5-6-Gal4* and *NB7-4-Gal4* lines are each  
192 expressed in a single neuroblast and its progeny during the Hb competence window and thus  
193 are ideal tools for expressing Dam or Dam:Hb in specific neuroblast lineages.

194 We next identified the early-born Hb+ progeny from both lineages, to ensure that each  
195 neuroblast lineage makes different Hb+ progeny. Dil clonal analyses show that both NB5-6 and  
196 NB7-4 make distinct populations of interneurons, but also similar populations of subperineurial  
197 glia, and their birth-order in the lineage has not been determined (Schmid et al., 1999; Schmidt et  
198 al., 1997). Therefore, we used *NB5-6-Gal4* to generate MultiColorFlipOut (MCFO; Nern et al.,  
199 2015) single neuron labelling among NB5-6 progeny. We repeatedly (n=31) identified a Hb<sup>+</sup>  
200 neuron that had a characteristic ipsilateral ascending projection, which we name the Chaise  
201 Lounge neuron due to its distinctive morphology; two segmentally repeated Chaise Lounge  
202 neurons are shown in Figure 2H; inset shows a Chaise Lounge neuron expressing Hb. We  
203 searched the EM reconstruction (Ohyama et al., 2015) and identified an identical Chaise Lounge  
204 neuron (Figure 2I). Thus, NB5-6 makes a distinctive ipsilateral neuron during its Hb expression

205 window. Similarly, we used *NB7-4-Gal4* to generate MCFO single cell labelling, but could not  
206 directly identify a Hb+ neuron either due to loss of Hb from early-born neurons prior neuronal  
207 differentiation, or due to lack of gal4 expression in these neurons. Instead, we used multiple  
208 criteria to identify a putative early-born neuron, the G neuron, using MARCM clones (Figure 2J),  
209 and EM reconstruction (Figure 2K). Our criteria for assigning this neuron as early-born include (i)  
210 presence of the neuron in full NB7-4 clones (Figure 2J) but not in the NB7-4-Gal4 pattern (Figure  
211 2, supplement 1), which misses early-born neurons; (ii) cell body position next to the neuropil,  
212 where most Hb+ neurons are located (Kambadur et al., 1998); and (iii) close morphological  
213 match to the early-born grasshopper G neuron, including ascending and descending projections  
214 in the most lateral connective tract (Raper et al., 1983). Finally, we note that all NB7-4 neuronal  
215 progeny have contralateral axons (Schmid et al., 1999; Schmidt et al., 1997), whereas the NB5-6  
216 early-born Chaise Lounge neuron has ipsilateral projections. Thus, we conclude that NB5-6 and  
217 NB7-4 produce different neurons during the Hb expression window. This makes NB5-6 and  
218 NB7-4 an appropriate model system to characterize how different spatial patterning cues  
219 produce distinct Hb+ early-born cell types.

### 221 **Generation of a functional, non-toxic Dam:Hb fusion protein**

222 The first step in using the TaDa method to map Hb occupancy in the NB5-6 and NB7-4 lineages  
223 is to generate a functional, non-toxic Dam:Hb fusion protein. Although other Dam constructs  
224 have been shown to be non-toxic (Aughey et al., 2018; Marshall et al., 2016; Southall et al.,  
225 2013), this is the first use of Dam:Hb and its toxicity is unknown. We used standard methods to  
226 generate a *UAS-LT3-Dam:hb* transgene where the first open reading frame (ORF) encodes  
227 Cherry and the second ORF encodes Dam:Hb (see Figure 1D,F); placing the Dam fusion protein  
228 in the second ORF is important to keep both Dam and Hb levels extremely low, which reduces  
229 toxicity and increases specificity of DNA binding (Southall et al., 2013).

230 To determine if Dam:Hb is toxic, we expressed the fusion protein throughout the nervous  
231 system (*sca-Gal4 UAS-Dam:Hb*) and ubiquitously (*Da-Gal4 UAS-Dam:Hb*), and observed no  
232 effect on embryonic viability (Figure 3A). To determine whether the Hb portion of the Dam:Hb  
233 fusion protein was functional, we assayed for its ability to generate ectopic Eve+ U neurons,  
234 despite being expressed at very low levels. In wild type, NB7-1 generates five Eve+ U neurons,  
235 including the Hb+ early-born U1 and U2 neurons, and extending neuroblast expression of Hb  
236 produces many ectopic Eve+ U1/U2 neurons (Isshiki et al., 2001; Pearson and Doe, 2003). We  
237 observed that expression of Dam:Hb was capable of inducing a small number of ectopic Eve+  
238 neurons (Figure 3B), despite the low levels of Dam:Hb, showing that Dam:Hb is functional. We  
239 conclude that Dam:Hb is non-toxic in embryos, and that it is functional for inducing early-born  
240 neuronal identity.

241 The fact that Dam:Hb can induce early-born neuronal identity suggests that it can bind the  
242 same genomic targets as Hb, but we wanted to determine this important point experimentally.  
243 The TaDa method involves comparing Dam genomic binding to Dam:Hb genomic binding, with a  
244 normalised ratio used to identify sites preferentially bound by the Dam:Hb fusion protein  
245 (Marshall and Brand, 2015; Southall et al., 2013). We expressed Dam or Dam:Hb in all cells  
246 throughout embryogenesis, measured the quantile normalised ratio between them to identify  
247 Dam:Hb binding sites (see methods), and performed three biological replicates at embryonic

248 stage 17. We found that the biological replicates showed high Pearson correlation coefficients  
249 (Figure 3C, left), and were qualitatively very similar along the entire fourth chromosome (Figure  
250 3C, right). Most importantly, we compared Dam:Hb genomic occupancy with published Hb  
251 genomic occupancy determined by chromatin immunoprecipitation (ChIP) (Bradley et al., 2010;  
252 Li et al., 2008). A comparison over 700 kb of genomic DNA on chromosome 3R showed  
253 qualitatively similar Dam:Hb and Hb ChIP binding profiles (Figure 3D). Indeed, enriched Dam:Hb  
254 binding was detected at eight of the nine known Hb target genes (Lyne et al., 2007) (Figure 3E,  
255 Figure 3-supplement 1). We next compared the similarities in Hb occupancy as reported by  
256 these two techniques at the genomic level. To do this, we ran the MACS2 peak caller (Zhang et  
257 al., 2008) on the two datasets and identified 6,597 and 6,656 regions significantly enriched for  
258 Dam:Hb and Hb ChIP respectively (see methods). We found that 1,972 regions were shared  
259 between the two (29.89% of ChIP peaks and 29.62% of Dam:Hb peaks). When broad peaks  
260 were used for this analysis, 2,394 regions were shared between the two, or 33.74% of ChIP  
261 peaks and 45.13% of Dam:Hb peaks; and when the narrow peaks were extended to 2kb on  
262 either side of the peak summit, 2,207 regions were shared between the two, or 57.53% of ChIP  
263 peaks and 60.37% of Dam:Hb peaks. A Monte Carlo analysis on the narrow peak overlap  
264 showed this was highly significant, detecting only 6.16% overlap with a set of random peaks  
265 (100 iterations,  $p$ -value  $< 1 e^{-300}$ , see methods). Correspondingly, we found high ChIP signals at  
266 the Dam:Hb binding sites and vice versa (Figure 3F,G, Figure 3-supplement 2). Importantly, this  
267 overlap in occupancy was not seen when the Dam:Hb data was compared with the ChIP-seq  
268 data of any other transcription factor, such as Ftz or Bcd (Figure 3G), demonstrating the  
269 specificity of the method. Additional support for the accuracy of Dam:Hb binding is that the  
270 known Hb DNA-binding motif is the most enriched motif at Dam:Hb binding sites (Figure 3 -  
271 supplement 3). Taken together, these results show that Dam:Hb binding closely mimics  
272 endogenous Hb binding.

273

### 274 **NB5-6 and NB7-4 lineages have different Hb-bound loci**

275 At this point we have validated two neuroblast-specific Gal4 lines, as well as shown that Dam:Hb  
276 genomic binding is both reproducible and matches published Hb ChIP data in stage 9 whole  
277 embryos. However, to test the two models of spatial and temporal integration we had to use  
278 Dam:Hb in the NB5-6 or NB7-4 lineages – much smaller pools of cells – to determine whether Hb  
279 genomic targets were the same or different in these spatially distinct NB lineages. Therefore, our  
280 next step was to determine if we could get reproducible Dam:Hb binding data from this small  
281 pool of cells, and with shorter Dam:Hb exposure than previously reported (Marshall and Brand,  
282 2017; Southall et al., 2013; Widmer et al., 2018). We expressed Dam:Hb in a single neuroblast  
283 lineage in each hemisegment (about 200 cells in the ~50,000 cell embryo) and for five hours  
284 (from embryonic stage 9-12). Previous experiments had expressed Dam constructs in a higher  
285 fraction of cells and for  $\geq 12$ h (Cheetham et al., 2018; Southall et al., 2013; Widmer et al., 2018).  
286 We expressed Dam:Hb using each of two neuroblast-specific Gal4 lines (*NB5-6-Gal4* and *NB7-  
287 4-Gal4*) and purified DNA from stage 12 embryos, near the end of the Hb competence window  
288 (see methods). We performed three biological replicates for each neuroblast and observed  
289 excellent reproducibility across all replicates (Figure 4A). We conclude that we can get a  
290 reproducible Dam:Hb signal from a single neuroblast lineage during the Hb competence window.

291 Next, we wanted to determine whether Dam:Hb binds the same or different loci in the two  
292 different neuroblasts. The high correlation between biological replicates for each neuroblast, plus  
293 the lack of correlation between the two neuroblasts, provided a gross indication that Dam:Hb  
294 has unique binding sites in each neuroblast lineage (Figure 4A). We expected the vast majority of  
295 loci in the genome to show a similar occupancy of Dam:Hb, because most genes are not  
296 predicted to regulate NB5-6/NB7-4 differences, and indeed, comparing Hb binding along the  
297 entire fourth chromosome shows qualitative similarities between the two NB lineages (Figure 4B).  
298 This is also evident at genes known to be expressed in and regulated by Hb across many  
299 neuroblast lineages – e.g. *Kr*, *pdm2* and *zfh2* (Isshiki et al., 2001) (Figure 4 – Supplement 1).  
300 These similarities confirm the reproducibility of Dam:Hb binding in two distinct neuroblast  
301 lineages.

302 To begin our analysis of differential Dam:Hb binding between NB5-6 lineage and NB7-4  
303 lineages, we first ran the MACS2 peak caller (Zhang et al., 2008) on the six datasets – three  
304 replicates of NB5-6 lineage and three replicates of NB7-4 lineage – to identify regions  
305 significantly bound by Hb in each sample. The rest of our analyses focussed on the significantly  
306 bound Hb loci in the two NB lineages. We used the R Bioconductor package DiffBind (Ross-  
307 Innes et al., 2012) to identify 4,224 differentially bound loci in the two NB lineages: 2,007 that  
308 were enriched for Dam:Hb binding in the NB5-6 lineage, and 2,217 that were enriched for  
309 Dam:Hb binding in the NB7-4 lineage (Figure 4C; Supplemental Table 1). In addition, there were  
310 2,860 loci occupied by Dam:Hb in both neuroblast lineages (Supplemental Table 1). Importantly,  
311 while the read densities at individual loci are similar between replicates, they are strikingly  
312 different between the two neuroblast lineages.

313 Next we represented the differentially bound loci using a volcano plot, where the magenta  
314 dots highlight the most significantly differential loci with more than 2-fold change and an FDR of  
315  $\leq 0.01$  (Figure 4D). This threshold corresponds to 718 Hb enriched loci in NB5-6 lineage and 504  
316 Hb enriched loci in NB7-4 lineage (Supplemental Table 1), which is what we use for all  
317 subsequent analyses. The genes closest to the top five differentially occupied loci in each  
318 neuroblast are marked in this plot, and shown in Figure 4E,F. Based on these results, we  
319 conclude that Dam:Hb binds different loci in different neuroblasts. This clearly rules out the  
320 independent specification model where Hb has identical binding sites in different neuroblasts.

321

### 322 **Different chromatin states in NB5-6 and NB7-4 lineages**

323 We next wanted to understand how STFs might influence TTF genomic binding. Given the order  
324 of their action – STFs acting early in the neuroectoderm, and TTFs acting later in the delaminated  
325 NB – one possibility is that STFs generate different open/closed chromatin landscapes in each  
326 neuroblast such that TTFs have access to different loci in each neuroblast. This would predict  
327 that spatially distinct NBs would have different open/closed chromatin landscapes. To determine  
328 if this were indeed true, we performed chromatin accessibility profiling by Dam only (CaTaDa),  
329 which exploits the ability of the Dam protein to bind open chromatin domains (Aughey et al.,  
330 2018). We first expressed Dam in all cells throughout embryogenesis using *Da-Gal4* and  
331 observed excellent reproducibility between biological replicates both qualitatively and  
332 quantitatively (Figure 5A, red tracks in C). We next wanted to confirm that Dam only binding in  
333 the embryo correlates with open chromatin domains, as has been shown in other cell types

334 (Aughey et al., 2018). To do this, we analysed the Dam only signal around the DNase I  
335 hypersensitive sites (peaks) made available by the BDTNP consortium (Thomas et al., 2011) and  
336 found enriched Dam signals around the DNase I peaks, as well as qualitative similarities between  
337 the two (Figure 5B, compare red and ochre tracks in C). We observed 6,708 Dam only peaks  
338 were aligned with DNase I hypersensitive peaks (44.6% of all Dam only peaks; 33.9% of all  
339 DNase I peaks). A Monte Carlo analysis showed this was highly significant, detecting only 18.14  
340 % overlap with a set of random peaks (100 iterations, p-value  $< 1 e^{-300}$ , see methods). These data  
341 suggest that Dam only can be used to detect open chromatin in embryos.

342 We next sought to determine whether Dam only could be used to assay open chromatin in  
343 small pools of cells over a short period of time – e.g. in NB5-6 and NB7-4 lineages at stage 12.  
344 We performed three biological replicates of Dam only for each neuroblast, and observed  
345 excellent reproducibility in all but one replicate, so we used the two best replicates henceforth  
346 (Figure 5D). The reproducibility of the method can also be observed in the similar Dam binding  
347 patterns seen at representative control genes that are equally expressed in NB5-6 and NB7-4  
348 lineages (e.g. *Kr*, *pdm2* and *zfh2*), or along a large stretch of chromosome 4 (Figure 5,  
349 Supplement 1).

350 Next, we investigated whether there were global differences in chromatin states between the  
351 two neuroblast lineages. To do this, we first determined regions of significantly open chromatin  
352 in the two neuroblast lineages by running the MACS2 peak caller (Zhang et al., 2008) on the four  
353 best replicates, which gave us a ‘peakset’ of significantly open chromatin in NB5-6 and NB7-4  
354 lineages. We used these regions of open chromatin in both NB5-6 and NB7-4 lineages to  
355 conduct a differential analysis using the DiffBind package (Ross-Innes et al., 2012) and identified  
356 a total of 8,740 Dam only differentially bound loci, including 3,656 loci in the NB5-6 lineage and  
357 5,084 loci in the NB7-4 lineage. These regions of differential chromatin accessibility have been  
358 represented as an ‘MA plot’ with the NB5-6 differential open chromatin loci at the top and the  
359 NB7-4 differential open chromatin loci at the bottom (Figure 5E). We conclude that there are  
360 global differences in the open chromatin landscape between the NB5-6 and NB7-4 lineages.

361  
362 **Neuroblast-specific Hb-bound loci correlate with neuroblast-specific open chromatin domains**  
363 Chromatin accessibility has been shown to be the strongest determinant of TF occupancy on the  
364 genome (Guertin et al., 2012; Kaplan et al., 2011 ; Li et al., 2008). We wanted to determine if  
365 Dam:Hb binding was similarly responsive to the state of the chromatin in the NB5-6 and NB7-4  
366 lineages. To do this, we took all Dam:Hb-bound loci – both those specific for each neuroblast as  
367 well as those shared by both neuroblasts – and queried the state of the chromatin at these loci in  
368 each NB lineage. We found that Dam:Hb-bound loci in the NB5-6 lineage were enriched for open  
369 chromatin in that lineage (Figure 6 – supplement 1A), and similarly, Dam:Hb-bound loci in the  
370 NB7-4 lineage were enriched for open chromatin in that lineage (Figure 6 – supplement 1B). This  
371 suggests that Dam:Hb binding is indeed correlated with chromatin accessibility domains in both  
372 NB lineages (Figure 6 – supplement 1C).

373 If Dam:Hb preferentially occupies regions of open chromatin, we reasoned that the  
374 differentially occupied Dam:Hb loci in each NB lineage (lineage-specific Hb loci) must be  
375 correlated with differentially open chromatin in that neuroblast lineage (lineage-specific open  
376 chromatin). Indeed, NB5-6-specific Dam:Hb loci showed a strong enrichment for open chromatin



377 (Figure 6A, blue lines); strikingly, these same loci had closed chromatin in NB7-4 (Figure 6A,  
378 green lines). Similarly, NB7-4-specific Dam:Hb loci showed strong enrichment for open  
379 chromatin (Figure 6B, green lines), while these same loci had closed chromatin in NB5-6 lineage  
380 (Figure 6B, blue lines). Corresponding to this, we found 364 peaks, or 50.76% of the differential  
381 Dam:Hb peaks in NB5-6 overlapped with differentially open chromatin peaks in that lineage; and  
382 164 peaks or 32.74% of the differential Dam:Hb peaks in NB7-4 overlapped with differentially  
383 open chromatin peaks in that lineage. A Monte Carlo analysis showed these overlaps to be  
384 highly significant, detecting 5.23% overlap with a set of random peaks in NB5-6 and 6.75% in  
385 NB 7-4 (100 iterations, p-value  $< 1 e^{-300}$  for NB 5-6 and  $8.9 e^{-133}$  for NB 7-4, see methods). As a  
386 control, we assayed loci bound by Dam:Hb in both neuroblast lineages and found that there was  
387 no difference between lineages in open chromatin at these sites (Figure 6C). We confirmed these  
388 findings at the top five differentially bound Dam:Hb loci in the two neuroblast lineages. All but  
389 two of these differentially bound loci were also identified in the differential chromatin analysis;  
390 even the two that were not picked up in the analysis (*sqz* and *mspo*) were qualitatively different  
391 between the two neuroblast lineages (Figure 6D,E). We conclude that neuroblast-specific  
392 Dam:Hb binding occurs within neuroblast-specific accessible chromatin domains. This  
393 correlation suggests that either Hb binds where chromatin is open, or that Hb binding opens  
394 chromatin. The latter model seems unlikely, because both NB5-6 and NB7-4 are exposed to Hb  
395 expression, yet each neuroblast has specific open chromatin domains (see Discussion). We favor  
396 a model in which STF s generate neuroblast-specific open chromatin domains, leading to  
397 neuroblast-specific Hb occupancy.

398

### 399 **The row 5 spatial transcription factor Gsb is enriched at open chromatin and Hb-bound** 400 **loci in NB5-6, but not NB7-4**

401 If spatial factors generate lineage-specific chromatin landscapes as the sequential specification  
402 model proposes, then it's likely that lineage-specific STF occupancy will correspond to lineage  
403 specific chromatin accessibility. Gsb is one of the best studied STF s in the embryonic VNC. It  
404 has been shown to be both necessary and sufficient to determine the identity of the row 5 NBs  
405 (Bhat, 1996; Skeath et al., 1995). Not only is Gsb a functionally validated STF, but Gsb ChIP-chip  
406 data from 0-12h embryos are publicly available (Bonneaud et al., 2017). As NB5-6 is a row 5 NB  
407 lineage specified by Gsb, it gave us the opportunity to test the sequential specification model  
408 more deeply. We asked whether Gsb occupancy was enriched at regions of accessible  
409 chromatin in the NB5-6 lineage. We plotted the Gsb ChIP-chip signal around all NB5-6 open  
410 chromatin loci and compared this with Gsb ChIP-chip signal around NB7-4 open chromatin loci.  
411 Indeed, we found an enrichment of Gsb signal specifically around NB5-6 open chromatin and  
412 not NB7-4 open chromatin (Figure 7A). A Monte Carlo analysis found this enrichment to be highly  
413 significant (average real NB5-6/ NB7-4 fold change = 2.198, average simulated NB5-6/ NB7-4  
414 fold change = 0.922, 100 random iterations, p-value =  $1.19119 e^{-62}$ ). This supports the hypothesis  
415 that lineage-specific STF s generate lineage-specific chromatin landscapes.

416 Finally, we reasoned that if Hb preferentially binds to regions of accessible chromatin, and  
417 STF occupancy correlates with open chromatin in a lineage-specific manner, then the lineage-  
418 specific Hb occupancy that we observe in NB5-6 should correlate with lineage specific STF  
419 occupancy. We therefore plotted Gsb signal around NB5-6-enriched Hb loci and found a

420 corresponding enrichment of Gsb occupancy at these regions (Figure 7B, blue line). In contrast,  
421 the NB7-4-enriched Hb loci did not show any such enrichment (Figure 7B, green line). A Monte  
422 Carlo analysis found this enrichment to be highly significant (average real NB5-6/NB7-4 fold  
423 change = 2.219, average simulated NB5-6/NB7-4 fold change = 1.088, 100 random iterations, p-  
424 value =  $5.1039 \times 10^{-9}$ ; see methods). We conclude that loci differentially bound by Hb in NB5-6 are  
425 enriched for Gsb occupancy, although we note that occupancy may occur at different times (Gsb  
426 earlier, Hb later).

427 Taken together, these data support the sequential specification model, where a transiently  
428 expressed STF (e.g. Gsb) sculpts a lineage-specific chromatin landscape in NB lineages (eg.  
429 NB5-6), this determines lineage-specific binding of TTFs (e.g. Hb), which can in turn specify  
430 different neural fates in different NB lineages (Figure 8).

431

## 432 Discussion

433

434 Since its first report, Targeted DamID has been used in multiple cell types, in both *Drosophila*  
435 and mammalian embryonic stem cells (ESCs), for mapping transcription factor binding  
436 (Cheetham et al., 2018), open chromatin domains (Aughey et al., 2018), chromatin states  
437 (Marshall and Brand, 2017), and for mapping paused or transcribed loci (Southall et al., 2013;  
438 Widmer et al., 2018). In all cases, the number of cells expressing the Dam constructs are  
439 relatively large: ~10,000 FACS purified ESCs (Cheetham et al., 2018) and ~5000 mushroom body  
440 neurons per brain (Widmer et al., 2018). In our study we analyze the smallest percentage of cells  
441 to date - we calculate that there are between 8-12 cells in each hemisegment expressing Dam  
442 constructs; with a total of 11 segments (three thoracic, eight abdominal) that would give a  
443 maximum of 264 cells per embryo, or about 0.5% of the estimated 50,000 cells per embryo.  
444 Furthermore, we pushed the limits of the technique by allowing just 5 hours of Dam or Dam:Hb  
445 expression. It's likely that this restrictive condition was successful in the case of a transcription  
446 factor-DNA interaction, which is stable during the time window; it might not be sufficient for  
447 factors such as RNA Pol II that require processivity through a gene. The ability to query  
448 transcription factor occupancy in such a precise manner – in a small subsets of cells over short  
449 periods of time – will encourage new uses of the method, such as studying the determination of  
450 cellular identities during development, upon reprogramming, or even in response to stimuli.

451 Using traditional methods of studying protein-DNA interactions, Hb targets in early  
452 embryogenesis have been well-characterized (Berman et al., 2002; Hoch et al., 1991; Rivera-  
453 Pomar et al., 1995; Struhl et al., 1992), yet little is known about Hb targets in the CNS, and  
454 nothing is known about neuroblast lineage-specific targets that specify lineage-specific neuronal  
455 identity. Here we've reported the first description of Hb occupancy *in vivo* within the genome of  
456 individual neuroblast lineages. Our study identified many loci that were similarly occupied in the  
457 two lineages, which are likely to consist of regulatory modules common to both lineages such as  
458 pan-neuronal specification or the progression of the temporal series. The latter example consists  
459 of Hb activating *Kr* and repressing *pdm2* in most neuroblast lineages. Indeed we find that Hb  
460 binds to both loci in NB5-6 and NB7-4 lineages, confirming previous observations that Hb  
461 directly represses *pdm2* and activates *Kr* in multiple neuroblast lineages (Kambadur et al., 1998;

462 Tran et al., 2010). Hb is also likely to directly repress *zfh2* in most neuroblast lineages (CQD,  
463 unpublished results) and our data show that the *zfh2* locus is indeed equivalently occupied in  
464 both neuroblast lineages. Apart from the commonly regulated loci, we identified over 100 loci  
465 that are differentially bound by Hb in NB5-6 or NB7-4. These are excellent candidates for  
466 lineage-specific neuronal specification.

467 NB5-6 and NB7-4 develop adjacent to each other during the earliest (S1) wave of neuroblast  
468 delamination. They share a common lateral Msh<sup>+</sup> spatial column, but are in different  
469 anterior/posterior spatial domains (NB5-6 is Gsb<sup>+</sup>, NB7-4 is En<sup>+</sup>). Although NB5-6 and NB7-4  
470 make different early-born neurons, they share a common ability to make subperineurial glia and  
471 neurons that project through the posterior commissure (Schmid et al., 1999; Schmidt et al.,  
472 1997). It is interesting to speculate that their common properties are due to their shared  
473 columnar spatial position, whereas their differences are due to different anterior/posterior spatial  
474 cues.

475 We show that ~1200 Hb-bound loci are different in NB5-6 and NB7-4 lineages, and that the  
476 chromatin at these sites is preferentially open, as determined by relatively high Dam binding. In  
477 some cases Dam:Hb occupancy is broader than Dam (open chromatin) occupancy; this could be  
478 due to Dam:Hb maintaining occupancy longer than Dam alone. The strong correlation between  
479 Dam:Hb binding and open chromatin could be due to Hb binding to previously opened  
480 chromatin domains, or Hb acting as a pioneer factor to open chromatin. We do not favor the  
481 latter mechanism because Hb binds some sites in NB5-6 but not in NB7-4 (and vice versa)  
482 showing that it is not sufficient to open chromatin. It is more likely that the STF Gsb acts as a  
483 pioneer to open chromatin in row 5 neuroblasts, similar to mammalian Pax family members  
484 (Budry et al., 2012; Mayran et al., 2015), and then subsequently Hb binding occurs within the  
485 available pool of open chromatin in each neuroblast. Similar findings have been reported for the  
486 estrogen receptor (ER), which is endogenously expressed in endometrial and breast cancer cell  
487 lines: its targets are different in both cell types, and the differential binding corresponds to  
488 differentially open chromatin (Gertz et al., 2013). Our conclusions are in agreement with studies  
489 showing that DNA accessibility, not cooperative or competitive interactions, have the strongest  
490 impact on transcription factor binding (Kaplan et al., 2011; Li et al., 2008). Similarly, this model is  
491 supported by in vitro protein-DNA studies that eliminate chromatin state contribution to these  
492 interactions (Guertin et al., 2012).

493 The specific enrichment of Gsb occupancy at regions of accessible chromatin in NB5-6 is a  
494 striking result that supports this hypothesis despite the Gsb ChIP data being from whole  
495 embryos. Ideally, similar experiments need to be conducted with Dam:Gsb in NB5-6 and Dam:En  
496 in NB7-4 lineage to determine correspondence of STF occupancy and chromatin accessibility,  
497 as well as STF and TTF occupancy in the NB lineages. The advantage of the *Drosophila* model is  
498 that these relationships can be rigorously tested. For example, mutational inactivation of the  
499 relevant STF, while assaying chromatin accessibility or Hb occupancy in a lineage-specific way  
500 will demonstrate the causal link between the STF and chromatin landscape, and STF and Hb  
501 occupancy. Similarly, targeting chromatin modifiers to select loci while assaying Hb occupancy  
502 will demonstrate a causal link between chromatin state and Hb occupancy. To definitively rule  
503 out the possibility that Hb acts as a pioneer in these lineages, it may be feasible to misexpress or

504 mutate Hb, to determine the effect on chromatin accessibility. These are technically difficult  
505 studies, beyond the scope of this paper.

506 Although we have provided evidence that Hb-bound loci are chosen from neuroblast-specific  
507 open chromatin domains, this does not rule out that sequential specification occurs via lineage-  
508 specific STFs/STF-target genes acting as Hb cofactors to bias Hb binding in each lineage.  
509 However, we have been unable to find any *de novo* DNA motif enriched within 1kb of Hb-bound  
510 genomic loci, either neuroblast-specific loci or within all Hb-bound loci. This is consistent with  
511 Hb acting independently, but we can't rule out the possibility of Hb acting with co-factors.

512 Our study, coming almost two decades after the first descriptions of spatial and temporal  
513 patterning in *Drosophila* neural stem cells (Isshiki et al., 2001), has for the first time explored the  
514 mechanism by which spatial and temporal factors could be integrated to generate neuroblast-  
515 specific neuronal progeny. Only recently has it been possible to probe TTF DNA-binding and  
516 chromatin landscapes within two distinct neuroblast lineages – due to the parallel advances in  
517 genetic tools, functional genomics, and our ability to manipulate the genome. Given the  
518 conservation of mechanisms in generating neural diversity in vertebrates and invertebrates, and  
519 exquisite ways in which the genome can now be manipulated in different organisms, it is now  
520 possible to determine if similar mechanisms generate diversity during vertebrate neurogenesis.

## 521 **Methods**

522

### 523 **Fly lines**

524 Fly stocks were obtained from the Bloomington Drosophila Stock Center (Bloomington, IN USA)  
525 and, unless otherwise stated, were grown on cornmeal media at 25°C. *UAS-LT3-Dam* flies were  
526 kindly provided by Andrea Brand, *19B03<sup>[ADJ]</sup>*; *18F07<sup>[DBD]</sup>* was a gift from Gerald Rubin, and *Lbe-(K)-*  
527 *Gal4* (called *NB5-6-Gal4* here) was a gift from Stephan Thor. To generate MCFO clones (Nern et  
528 al., 2015) with *NB5-6-Gal4* or *NB7-4-Gal4*, we crossed *hsFLP*; *UAS-MCFO* females to Gal4 line  
529 males. 0-1 hour eggs were collected, aged at 25°C until stage 8 and given a 37°C heat shock for  
530 20 minutes then aged at 25°C or 18°C until stage 17. We used MARCM (Lee and Luo, 1999) with  
531 *engrailed-Gal4* to generate NB7-4 clones, which were unambiguously identified by the presence  
532 of channel glia (Schmid et al., 1999; Schmidt et al., 1997).

533

### 534 **Immunohistochemistry and confocal imaging**

535 Embryos were dechorionated in bleach for 3 minutes and fixed in 1:1::4% PFA:Heptane for 20-30  
536 minutes. Vitelline membranes were removed by shaking them vigorously in  
537 1:1::heptane:methanol. They were washed with blocking solution (1× PBS with 0.3% TritonX and  
538 0.1% BSA) for an hour. Primary antibodies were diluted in blocking solution. The samples were  
539 incubated on horizontal shaker at 4°C for 24 hrs after which they were washed with 0.3% PTX (1×  
540 PBS with 0.3% TritonX) and secondary antibody diluted in 0.3% PTX was added. The samples  
541 were incubated at 4°C overnight, washed 0.3% PTX, allowed to settle in 30% glycerol, then  
542 allowed to clear in 90% glycerol infused with Vectashield overnight. Primary antibodies used  
543 were: chicken anti-GFP (1:1000, abcam ab13970), mouse anti-engrailed (1:50, 4D9 DSHB); rat  
544 anti-gooseberry (1:10 of equal mix of 10E10 and 16F2, Holmgren Lab), rabbit anti-Hunchback  
545 (1:400), rabbit anti-Dan (1:1000), mouse anti mCherry (1:500, Clontech 632543), rabbit anti-  
546 V5::549 (1:400, Rockland 600-442-378), mouse anti-HA::488 (1:200, Cell signaling 2350S), rat  
547 anti-Ollas::650 (1:200, Novus NBP1-06713) and rabbit anti-Eve (1:500). All samples were imaged  
548 on ZeissLSM700 or ZeissLSM710 confocal microscope. Optical sections were acquired at 0.75  
549 µm intervals with a picture size of 1024 × 1024 pixels. Images were processed in the open source  
550 software FIJI (<http://fiji.sc>).

551

### 552 **Generation of Dam:Hb**

553 To generate *UAS-LT3-Dam:hb*, full-length *hb* CDS was PCR amplified from BACR01F13 and  
554 cloned into *pUAST-attB-LT3-NDam* (a gift from Andrea Brand) using NotI and XbaI sites to fuse  
555 Dam to the N-terminus of Hb. As spontaneous mutations are known to arise in the Dam sequence  
556 upon transformation (Marshall et al., 2016), its sequence integrity was tested at each  
557 transformation step, and prior to injections, all three elements - Dam, Hb and Cherry sequences  
558 were confirmed to be preserved. Transgenic flies with the construct integrated at the attP2 landing  
559 site were generated by BestGene Inc.

560

### 561 **Dam:Hb and Dam genomic binding**

562 For verifying the Dam:Hb flies, about 1500 females of *UAS-LT3-Dam* and *UAS-LT3-Dam:hb* flies  
563 were crossed to about 500 males of *Da-Gal4* in egg collection cages placed at 25°C. Embryos  
564 were collected every two hours and aged for 16 hours at 25°C, then dechorionated with bleach to



565 avoid contaminants, washed thoroughly with de-ionized water and preserved at -20°C until  
566 sufficient material was collected - for each replicate, 50mg of control and experimental embryos.  
567 For stage 12 neuroblast TaDa experiments, about 5,000-6,000 *UAS-LT3-Dam* and *UAS-LT3-*  
568 *Dam:hb* flies were crossed to about 3,000 *Lbe-K-Gal4* or *19B03<sup>[ADJ]</sup>/18F07<sup>[DBD]</sup>* flies. Embryos were  
569 collected every two hours and aged for 7.5 hours at 25°C, and similarly treated until sufficient  
570 material was collected - for each replicate, 4 X 1.5 µL tubes of 50 mg of control and experimental  
571 embryos.

572 The TaDa experimental pipeline was followed according to (Marshall et al., 2016), with a few  
573 alterations to optimize for small cell numbers and short duration of Dam expression. Briefly, the 4  
574 tubes of each replicate were thawed on ice, processed separately and in parallel until the PCR  
575 purification step after the DpnI digestion step; subsequently, an additional PCR purification step  
576 using standard Qiagen PCR purification columns was used to concentrate the DpnI digested  
577 product to 32 µL. Embryos were homogenized with an electric pestle and gDNA was extracted  
578 using the DNA Micro Kit (Qiagen, cat. no. 56304). Extreme care was taken to ensure that the  
579 gDNA remained intact – this was done by using wide bore tips to avoid fragmenting the DNA,  
580 pipetting deliberately, and avoiding any rough shaking/tipping. gDNA was digested with DpnI for  
581 14-16 hours in a thermocycler then PCR purified. MyTaq HS DNA polymerase kit (Bioline, cat. no.  
582 BIO-21112; not the Advantage 2 cDNA polymerase from Clontech) was used for amplification  
583 and 21 PCR cycles we used. Sequencing libraries were prepared according to the Illumina  
584 TruSeq DNA library protocol. The samples were sequenced on the Illumina HiSeq4000 at 100  
585 base pairs and about 20-60 million single end reads per sample.

## 586 **Bioinformatic Analysis**

587 **Quality control.** Each file was assessed for quality using FastQC (Andrews, 2010). Reads with  
588 quality score less than 30 were discarded. Any contaminants were removed using BBSplit of the  
589 BBmap suite (<https://sourceforge.net/projects/bbmap/> ).

592 **The damidseq pipeline** was used to generate log2 ratio files (Dam:hb/Dam) in GATC resolution  
593 as described previously (Marshall and Brand, 2015). Briefly, the pipeline uses Bowtie2  
594 (Langmead and Salzberg, 2012) to align reads to dm6, the reads are extended to 300bp (or to the  
595 closest GATC, whichever is first) and this .bam output is used to generate the ratio file  
596 (.bedgraph). The bedgraph files were used for data visualization on IGV 2.4.1 (Robinson et al.,  
597 2011; Thorvaldsdóttir et al., 2013) and the read extended bam files were used for peak calling.

599 **Correlation coefficients** between biological replicates for Da-Gal4 Hb TaDa and Da-Gal4 CaTaDa  
600 were computed using the multiBamSummary and plotCorrelation functions of DeepTools. For  
601 NB5-6 and NB7-4 Hb TaDa and CaTaDa, where differential analyses were conducted, the  
602 correlation coefficients computed by DiffBind (Ross-Innes et al., 2012) are represented.

604 **Peak calling.** For **TaDa** experiments, MACS2 (v2.1.1) (Zhang et al., 2008) was used to call narrow  
605 peaks on sorted, read extended bam files of Dam:Hb, with a single merged Dam only as a control  
606 provided for each replicate. MACS2 (v2.1.1) was also used to call peaks on Hb ChIP-seq data.  
607 For this, dm3 aligned Hb ChIP-seq and input files (in bowtie output format) were downloaded from  
608

609 NCBI (GEO accession number GSE20369; HB2) and converted to sam format using  
610 bowtie2sam.pl from the SAMtools suite. These were converted to bam and CrossMap (Zhao et  
611 al., 2014) was then used to liftOver both the input and Hb files from dm3->dm6. deepTools was  
612 used to generate the ratio files for subsequent analyses. For **CaTaDa** experiments, narrow peaks  
613 were called on sorted, read extended bam files of Dam only using MACS2 (v2.1.1) without  
614 controls.

615  
616 Peak overlap. bedtools intersect was used for computing peak overlaps. An overlap of 1 basepair  
617 or more was considered an overlap. *Hb ChIP-seq vs. Hb TaDa*: narrow peak output from MACS2  
618 were used for both files. *Da-Gal4 CaTaDa vs. DNaseI*: the MACS2 generated narrow peaks for  
619 *Da-Gal4* CaTaDa was supplied along with the stage 11 DNaseI peak file, which was downloaded  
620 from BDTNP and lifted over from dm2->dm6 using CrossMap. *Differential Hb vs. differential*  
621 *chromatin*: the differentially bound sites identified by DiffBind (Ross-Innes et al., 2012) were saved  
622 as bed files and provided to bedtools intersect to assess overlap percentage.

623  
624 Monte Carlo analysis. To check for the significance of peak/signal overlap, a Monte Carlo analysis  
625 was performed. *Hb TaDa vs. Hb ChIP*: Hb ChIP was taken as the reference, and an equal  
626 number of random peaks were generated such that the number and length of peaks for each  
627 chromosome remained the same. These random peaks were used to check for overlap with Hb  
628 TaDa. A 100 such iterations were performed, and an average overlap calculated for the random  
629 overlap. Z-score and p-value was calculated between the average random overlap and the actual  
630 overlap. A custom written script was used to perform this analysis (Aughey et al., 2018). *Da-Gal4*  
631 *CaTaDa vs. DNaseI*: Similar analysis as above was used with DNaseI as the reference.  
632 *Differential Hb and Differential chromatin*: Differentially bound, thresholded Hb peaks of NB5-6  
633 and NB7-4 were taken as the reference and an equal number of random peaks were generated  
634 such that the number and length of peaks for each chromosome remained the same. These  
635 random peaks were used to check for overlap with the differentially bound chromatin loci in the  
636 respective NB. A 100 such iterations were performed, and an average overlap calculated for the  
637 random overlap. The Z-score and p-value were calculated between the average random overlap  
638 and the actual overlap. *Hb Gsb signal at 5-6 and 7-4 chromatin and enriched Hb loci*: 'bedtools  
639 slop' was used to extend the 5-6 and 7-4 peaksets to 4kb (2kb on either side of the peak center).  
640 An equal number of random peaks were generated for 5-6 and 7-4 as in the actual data  
641 (respecting distribution of peaks on the chromosomes). 'bedtools shuffle' was used to generate  
642 these random peaks. The Gsb data obtained from Florence Maschat was converted from wig to  
643 bedgraph using 'wig2bed' from bedops, then dm3->dm6 using CrossMap, and finally from  
644 bedgraph to bigwig using 'bedGraphToBigWig' from kentUtils ([https://github.com/ENCODE-](https://github.com/ENCODE-DCC/kentUtils)  
645 [DCC/kentUtils](https://github.com/ENCODE-DCC/kentUtils)). 'bigWigAverageOverBed' from kentUtils was used to generate the average Gsb  
646 signal at each peak. The average signal for each iteration was generated using awk. The  
647 difference in average Gsb signal between (randomly generated) NB5-6 and (randomly generated)  
648 NB7-4 was calculated for a 100 such iterations. The difference between average Gsb signal for  
649 the real data (i.e. 5-6 enriched Hb loci minus 7-4 enriched Hb loci) was similarly calculated. Z  
650 scores and p-values were calculated based on these 100 simulations and real differences in Gsb  
651 signal. A bash script was written to automate the above steps (available upon request). Similar  
652 pipeline was used for comparisons with *bcd*, *kni*, *cad* and *Kr*.

653  
654 TaDa/CaTaDa signal comparisons with other data. The computeMatrix tool from deepTools was  
655 used to plot the signal distribution relative to reference points in [Figures 3F,G; 5B; 6A-C; 7A,B;](#)  
656 [and Figure 6, supplement 1](#). In all cases, signal files (of ChIP or TaDa data) were supplied as  
657 bigwig files, and peaks regions were supplied as bed files. [Figure 3F](#) peak file was the narrow  
658 peaks generated by MACS2 in the three Da-Gal4 Hb TaDa experiments; the Hb ChIP-seq ratio  
659 file was used as the signal file (see under peak calling for details). [Figure 3G](#) peak files for Hb,  
660 Bcd and Ftz were downloaded from BDTNP and were lifted-over from dm3->dm6 using  
661 CrossMap; the Hb TaDa signal was converted to bigwig using 'bedGraphToBigWig' from kentUtils  
662 (<https://github.com/ENCODE-DCC/kentUtils>). [Figure 5B](#) peak file was downloaded from BDTNP and  
663 was lifted-over from dm2->dm6 using CrossMap; the Da-Gal4 CaTaDa signal was converted to  
664 bigwig using 'bedGraphToBigWig' from kentUtils. [Figure 6A-C](#): separate region files were made  
665 from the DiffBind (Ross-Innes et al., 2012) output for NB5-6 enriched, 7-4 enriched and 'Not-  
666 Differentially Bound' Hb loci; NB5-6 and NB7-4 CaTaDa files were converted to bigwig using  
667 'bamCoverage' of deepTools. [Figure 6-supplement 1A,B](#): MACS2 generated narrow peaks for  
668 NB5-6 and NB7-4 were used; NB5-6 and NB7-4 CaTaDa files were converted to bigwig using  
669 'bamCoverage' of deepTools. [Figure 7A](#): All MACS2 generated narrow peaks on the NB5-6 and  
670 NB7-4 CaTaDa were supplied as the regions of open chromatin; Gsb ChIP-chip signal file was  
671 used (see under Monte Carlo analysis for details). [Figure 7B](#): separate region files were made  
672 from the DiffBind (Ross-Innes et al., 2012) output for NB5-6 enriched and 7-4 enriched Hb loci;  
673 Gsb ChIP-chip signal file was used (see under Monte Carlo analysis for details).

674  
675 Motif calling was performed using the findMotifs.pl tool from the Homer suite of tools. The top  
676 1,000 narrow peaks from MACS2 were supplied to Homer and *de novo* motif calling was  
677 performed on 300 kb on either side of the peak centre. Approximately 6.5 times the number of  
678 supplied peaks were used as background to calculate enrichment. Using all peaks gave  
679 comparable results, with Hb as the most enriched motif over background.

680  
681 Differential analyses in [Figures 4 and 5](#) were performed using DiffBind (Ross-Innes et al., 2012).  
682 Briefly, narrow peak output files were provided for each of the three replicates of NB5-6 and NB7-  
683 4, along with their aligned Dam:Hb ([Figure 4](#)) or Dam alone ([Figure 5](#)) bam files. An initial  
684 correlation was calculated between the samples (both between replicates and across NBs) at  
685 these loci. The number of overlapping reads at each region was calculated, normalized, and  
686 represented as a *binding affinity matrix*. This matrix data was used for the further differential  
687 binding analysis and assignment of FDR and p-values, which can be conducted using either  
688 DeSeq2 or edgeR packages. Data shown here are results from DeSeq2 based differential  
689 analyses. Correlation heatmap, binding affinity matrix, MA plots and volcano plots represented in  
690 [Figures 4 and 5](#) were generated using Diffbind (Ross-Innes et al., 2012).

## 691 **Acknowledgements**

692 We thank Keiko Hirono and Dylan Heussman for generating the Dam:Hb transgene; Sen-Lin Lai  
693 for [Figure 2J](#); Keiko Hirono for contributing to [Figure 3A](#); Andrea Brand for TaDa reagents;  
694 Stephan Thor for Lbe reagents; Gerry Rubin for 7-4-Gal4; Jan Trout for Figure 1 illustrations; and  
695 Maggie Weitzman and Douglas Turnbull at the UO Genomics facility. We thank Sen-Lin Lai,  
696

697 Brandon Mark, Heinrich Reichert, Vishaka Datta, Gabriel Aughey, and Richard Mann for  
698 comments on the manuscript. Stocks obtained from the Bloomington Drosophila Stock Center  
699 (NIH P40OD018537) were used in this study. Funding was provided by the Fulbright-Nehru  
700 Postdoctoral fellowship (SQS), HHMI (CQD, SQS, SC), and NIH HD27056 (CQD).

701

#### 702 **Author Contributions**

703 CQD and SQS conceived of the project; SQS generated data; TDS, SC and SQS performed  
704 bioinformatics, and all authors commented and approved of the manuscript.

705

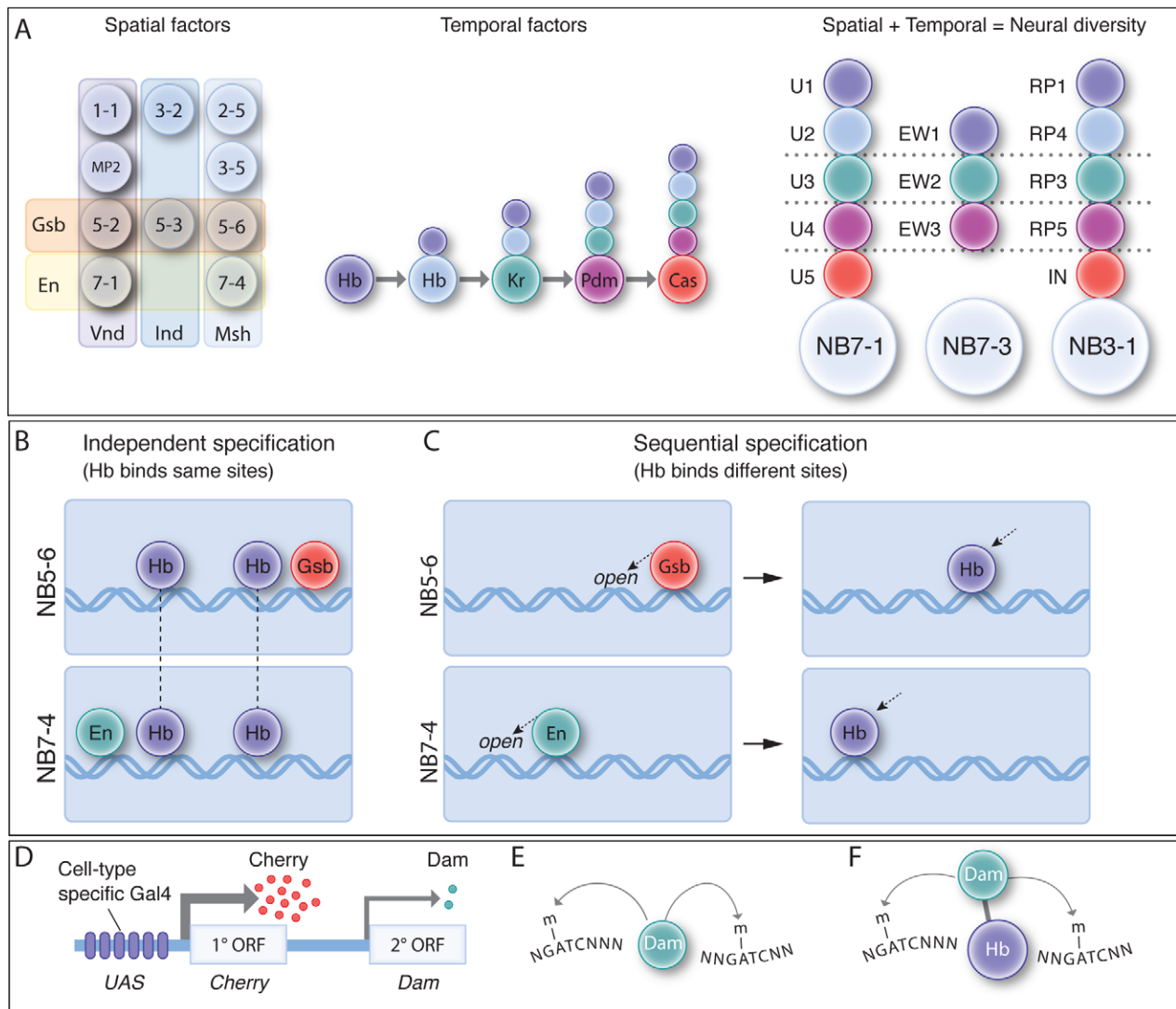
## 706 References

- 707 Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data.  
708 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 709 Aughey, G.N., Estacio Gomez, A., Thomson, J., Yin, H., Southall, T.D., 2018. CATaDa reveals global  
710 remodelling of chromatin accessibility during stem cell differentiation in vivo. **Elife** 7.
- 711 Baumgardt, M., Karlsson, D., Terriente, J., Diaz-Benjumea, F.J., Thor, S., 2009. Neuronal subtype  
712 specification within a lineage by opposing temporal feed-forward loops. **Cell** 139, 969-982.
- 713 Berman, B.P., Nibu, Y., Pfeiffer, B.D., Tomancak, P., Celniker, S.E., Levine, M., Rubin, G.M., Eisen, M.B.,  
714 2002. Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in  
715 pattern formation in the Drosophila genome. **PNAS** 99, 757-762.
- 716 Bhat, K.M., 1996. The patched signaling pathway mediates repression of gooseberry allowing neuroblast  
717 specification by wingless during Drosophila neurogenesis. **Development** 122, 2921-2932.
- 718 Bonneaud, N., Layalle, S., Colomb, S., Jourdan, C., Ghysen, A., Severac, D., Dantec, C., Negre, N.,  
719 Maschat, F., 2017. Control of nerve cord formation by Engrailed and Gooseberry-Neuro: A multi-step,  
720 coordinated process. **Dev Biol** 432, 273-285.
- 721 Bradley, R.K., Li, X.Y., Trapnell, C., Davidson, S., Pachter, L., Chu, H.C., Tonkin, L.A., Biggin, M.D., Eisen,  
722 M.B., 2010. Binding site turnover produces pervasive quantitative changes in transcription factor binding  
723 between closely related Drosophila species. **PLoS Biol** 8, e1000343.
- 724 Budry, L., Balsalobre, A., Gauthier, Y., Khetchoumian, K., L'Honore, A., Vallette, S., Brue, T., Figarella-  
725 Branger, D., Meij, B., Drouin, J., 2012. The selector gene Pax7 dictates alternate pituitary cell fates through  
726 its pioneer action on chromatin remodeling. **Genes Dev** 26, 2299-2310.
- 727 Cheetham, S.W., Gruhn, W.H., van den Aamele, J., Krautz, R., Southall, T.D., Kobayashi, T., Surani, M.A.,  
728 Brand, A.H., 2018. Targeted DamID reveals differential binding of mammalian pluripotency factors.  
729 **Development** 145, dev170209.
- 730 Chu-LaGraff, Q., Doe, C.Q., 1993. Neuroblast specification and formation regulated by wingless in the  
731 Drosophila CNS. **Science** 261, 1594-1597.
- 732 Deshpande, N., Dittrich, R., Technau, G.M., Urban, J., 2001. Successive specification of Drosophila  
733 neuroblasts NB 6-4 and NB 7-3 depends on interaction of the segment polarity genes wingless, gooseberry  
734 and naked cuticle. **Development** 128, 3253-3261.
- 735 Doe, C.Q., 2017. Temporal Patterning in the Drosophila CNS. **Annu Rev Cell Develop Biol** 33, 219-240.
- 736 Erclik, T., Li, X., Courgeon, M., Bertet, C., Chen, Z., Baumert, R., Ng, J., Koo, C., Arain, U., Behnia, R., del  
737 Valle Rodriguez, A., Senderowicz, L., Negre, N., White, K.P., Desplan, C., 2017. Integration of temporal and  
738 spatial patterning generates neural diversity. **Nature** 541, 365-370.
- 739 Gertz, J., Savic, D., Varley, K.E., Partridge, E.C., Safi, A., Jain, P., Cooper, G.M., Reddy, T.E., Crawford, G.E.,  
740 Myers, R.M., 2013. Distinct properties of cell-type-specific and shared transcription factor binding sites.  
741 **Molec Cell** 52, 25-36.
- 742 Grosskortenhaus, R., Robinson, K.J., Doe, C.Q., 2006. Pdm and Castor specify late-born motor neuron  
743 identity in the NB7-1 lineage. **Genes Dev** 20, 2618-2627.
- 744 Guertin, M.J., Martins, A.L., Siepel, A., Lis, J.T., 2012. Accurate prediction of inducible transcription factor  
745 binding intensities in vivo. **PLoS Genet** 8, e1002610.
- 746 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., Glass,  
747 C.K., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements  
748 required for macrophage and B cell identities. **Molec Cell** 38, 576-589.
- 749 Hoch, M., Seifert, E., Jackle, H., 1991. Gene expression mediated by cis-acting sequences of the Kruppel  
750 gene in response to the Drosophila morphogens bicoid and hunchback. **The EMBO journal** 10, 2267-2278.
- 751 Isshiki, T., Pearson, B., Holbrook, S., Doe, C.Q., 2001. Drosophila neuroblasts sequentially express  
752 transcription factors which specify the temporal identity of their neuronal progeny. **Cell** 106, 511-521.
- 753 Isshiki, T., Takeichi, M., Nose, A., 1997. The role of the msh homeobox gene during Drosophila  
754 neurogenesis: implication for the dorsoventral specification of the neuroectoderm. **Development** 124, 3099-  
755 3109.
- 756 Jessell, T.M., 2000. Neuronal specification in the spinal cord: inductive signals and transcriptional codes.  
757 **Nat Rev Genet** 1, 20-29.



- 758 Kambadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S.J., Odenwald, W.F., 1998. Regulation of POU  
759 genes by castor and hunchback establishes layered compartments in the Drosophila CNS. **Genes Dev** 12,  
760 246-260.
- 761 Kanai, M.I., Okabe, M., Hiromi, Y., 2005. seven-up Controls switching of transcription factors that specify  
762 temporal identities of Drosophila neuroblasts. **Dev Cell** 8, 203-213.
- 763 Kaplan, T., Li, X.Y., Sabo, P.J., Thomas, S., Stamatoyannopoulos, J.A., Biggin, M.D., Eisen, M.B., 2011.  
764 Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding  
765 during early Drosophila development. **PLoS Genet** 7, e1001290.
- 766 Kohwi, M., Doe, C.Q., 2013. Temporal fate specification and neural progenitor competence during  
767 development. **Nature Rev Neurosci** 14, 823-838.
- 768 Kohwi, M., Lupton, J.R., Lai, S.L., Miller, M.R., Doe, C.Q., 2013. Developmentally regulated subnuclear  
769 genome reorganization restricts neural progenitor competence in Drosophila. **Cell** 152, 97-108.
- 770 Lacin, H., Truman, J.W., 2016. Lineage mapping identifies molecular and architectural similarities between  
771 the larval and adult Drosophila central nervous system. **Elife** 5, e13399.
- 772 Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. **Nature Methods** 9, 357.
- 773 Lee, T., Luo, L., 1999. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal  
774 morphogenesis. **Neuron** 22, 451-461.
- 775 Li, X., Chen, Z., Desplan, C., 2013. Temporal patterning of neural progenitors in Drosophila. **Current Top**  
776 **Dev Biol** 105, 69-96.
- 777 Li, X.Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D.A., Iyer, V.N., Hechmer, A., Simirenko, L., Stapleton,  
778 M., Luengo Hendriks, C.L., Chu, H.C., Ogawa, N., Inwood, W., Sementchenko, V., Beaton, A., Weiszmann,  
779 R., Celniker, S.E., Knowles, D.W., Gingeras, T., Speed, T.P., Eisen, M.B., Biggin, M.D., 2008. Transcription  
780 factors bind thousands of active and inactive regions in the Drosophila blastoderm. **PLoS Biol** 6, e27.
- 781 Lyne, R., Smith, R., Rutherford, K., Wakeling, M., Varley, A., Guillier, F., Janssens, H., Ji, W., McLaren, P.,  
782 North, P., Rana, D., Riley, T., Sullivan, J., Watkins, X., Woodbridge, M., Lilley, K., Russell, S., Ashburner, M.,  
783 Mizuguchi, K., Micklem, G., 2007. FlyMine: an integrated database for Drosophila and Anopheles genomics.  
784 **Genome Biol** 8, R129.
- 785 Marshall, O.J., Brand, A.H., 2015. damidseq\_pipeline: an automated pipeline for processing DamID  
786 sequencing datasets. **Bioinformatics (Oxford, England)** 31, 3371-3373.
- 787 Marshall, O.J., Brand, A.H., 2017. Chromatin state changes during neural development revealed by in vivo  
788 cell-type specific profiling. **Nature Comm** 8, 2271.
- 789 Marshall, O.J., Southall, T.D., Cheetham, S.W., Brand, A.H., 2016. Cell-type-specific profiling of protein-DNA  
790 interactions without cell isolation using targeted DamID with next-generation sequencing. **Nat Protocols** 11,  
791 1586-1598.
- 792 Mayran, A., Pelletier, A., Drouin, J., 2015. Pax factors in transcription and epigenetic remodelling. **Seminars**  
793 **in Cell & Developmental Biology** 44, 135-144.
- 794 McDonald, J.A., Doe, C.Q., 1997. Establishing neuroblast-specific gene expression in the Drosophila CNS:  
795 huckebein is activated by Wingless and Hedgehog and repressed by Engrailed and Gooseberry.  
796 **Development** 124, 1079-1087.
- 797 McDonald, J.A., Holbrook, S., Isshiki, T., Weiss, J., Doe, C.Q., Mellerick, D.M., 1998. Dorsoventral patterning  
798 in the Drosophila central nervous system: the vnd homeobox gene specifies ventral column identity. **Genes**  
799 **Dev** 12, 3603-3612.
- 800 Miller, M.R., Robinson, K.J., Cleary, M.D., Doe, C.Q., 2009. TU-tagging: cell type-specific RNA isolation from  
801 intact complex tissues. **Nat Methods** 6, 439-441.
- 802 Nern, A., Pfeiffer, B.D., Rubin, G.M., 2015. Optimized tools for multicolor stochastic labeling reveal diverse  
803 stereotyped cell arrangements in the fly visual system. **PNAS** 112, E2967-2976.
- 804 Novotny, T., Eiselt, R., Urban, J., 2002. Hunchback is required for the specification of the early sublineage of  
805 neuroblast 7-3 in the Drosophila central nervous system. **Development** 129, 1027-1036.
- 806 Ohyama, T., Schneider-Mizell, C.M., Fetter, R.D., Aleman, J.V., Franconville, R., Rivera-Alba, M., Mensh,  
807 B.D., Branson, K.M., Simpson, J.H., Truman, J.W., Cardona, A., Zlatic, M., 2015. A multilevel multimodal  
808 circuit enhances action selection in Drosophila. **Nature** 520, 633-639.
- 809 Pearson, B.J., Doe, C.Q., 2003. Regulation of neuroblast competence in Drosophila. **Nature** 425, 624-628.

- 810 Prokop, A., Technau, G.M., 1994. Early tagma-specific commitment of *Drosophila* CNS progenitor NB1-1.  
811 **Development** 120, 2567-2578.
- 812 Raper, J.A., Bastiani, M., Goodman, C.S., 1983. Pathfinding by neuronal growth cones in grasshopper  
813 embryos. I. Divergent choices made by the growth cones of sibling neurons. **J Neuroscience** 3, 20-30.
- 814 Rivera-Pomar, R., Lu, X., Perrimon, N., Taubert, H., Jackle, H., 1995. Activation of posterior gap gene  
815 expression in the *Drosophila* blastoderm. **Nature** 376, 253-256.
- 816 Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011.  
817 Integrative genomics viewer. **Nature Biotechnology** 29, 24.
- 818 Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis,  
819 O., Ellis, I.O., Green, A.R., Ali, S., Chin, S.F., Palmieri, C., Caldas, C., Carroll, J.S., 2012. Differential  
820 oestrogen receptor binding is associated with clinical outcome in breast cancer. **Nature** 481, 389-393.
- 821 Schmid, A., Chiba, A., Doe, C.Q., 1999. Clonal analysis of *Drosophila* embryonic neuroblasts: neural cell  
822 types, axon projections and muscle targets. **Development** 126, 4653-4689.
- 823 Schmidt, H., Rickert, C., Bossing, T., Vef, O., Urban, J., Technau, G.M., 1997. The embryonic central  
824 nervous system lineages of *Drosophila melanogaster*. II. Neuroblast lineages derived from the dorsal part of  
825 the neuroectoderm. **Dev Biol** 189, 186-204.
- 826 Skeath, J.B., Zhang, Y., Holmgren, R., Carroll, S.B., Doe, C.Q., 1995. Specification of neuroblast identity in  
827 the *Drosophila* embryonic central nervous system by gooseberry-distal. **Nature** 376, 427-430.
- 828 Southall, T.D., Gold, K.S., Egger, B., Davidson, C.M., Caygill, E.E., Marshall, O.J., Brand, A.H., 2013. Cell-  
829 type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II  
830 occupancy in neural stem cells. **Develop Cell** 26, 101-112.
- 831 Stanojevic, D., Hoey, T., Levine, M., 1989. Sequence-specific DNA-binding activities of the gap proteins  
832 encoded by hunchback and Kruppel in *Drosophila*. **Nature** 341, 331-335.
- 833 Struhl, G., Johnston, P., Lawrence, P.A., 1992. Control of *Drosophila* body pattern by the hunchback  
834 morphogen gradient. **Cell** 69, 237-249.
- 835 Thomas, S., Li, X.Y., Sabo, P.J., Sandstrom, R., Thurman, R.E., Canfield, T.K., Giste, E., Fisher, W.,  
836 Hammonds, A., Celniker, S.E., Biggin, M.D., Stamatoyannopoulos, J.A., 2011. Dynamic reprogramming of  
837 chromatin accessibility during *Drosophila* embryo development. **Genome Biol** 12, R43.
- 838 Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-  
839 performance genomics data visualization and exploration. **Briefings in Bioinformatics** 14, 178-192.
- 840 Tran, K.D., Doe, C.Q., 2008. Pdm and Castor close successive temporal identity windows in the NB3-1  
841 lineage. **Development** 135, 3491-3499.
- 842 Tran, K.D., Miller, M.R., Doe, C.Q., 2010. Recombineering Hunchback identifies two conserved domains  
843 required to maintain neuroblast competence and specify early-born neuronal identity. **Development** 137,  
844 1421-1430.
- 845 Urbach, R., Technau, G.M., 2003. Molecular markers for identified neuroblasts in the developing brain of  
846 *Drosophila*. **Development** 130, 3621-3637.
- 847 Weiss, J.B., Von Ohlen, T., Mellerick, D.M., Dressler, G., Doe, C.Q., Scott, M.P., 1998. Dorsoventral  
848 patterning in the *Drosophila* central nervous system: the intermediate neuroblasts defective homeobox gene  
849 specifies intermediate column identity. **Genes Dev** 12, 3591-3602.
- 850 Widmer, Y.F., Bilican, A., Bruggmann, R., Sprecher, S.G., 2018. Regulators of Long-Term Memory Revealed  
851 by Mushroom Body-Specific Gene Expression Profiling in *Drosophila melanogaster*. **Genetics** 209, 1167-  
852 1181.
- 853 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M.,  
854 Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). **Genome Biol** 9, R137.
- 855 Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.P., Wang, L., 2014. CrossMap: a versatile tool for  
856 coordinate conversion between genome assemblies. **Bioinformatics (Oxford, England)** 30, 1006-1007.
- 857  
858



859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875

**Figure 1. Spatial and temporal cues are integrated to generate neuronal diversity**

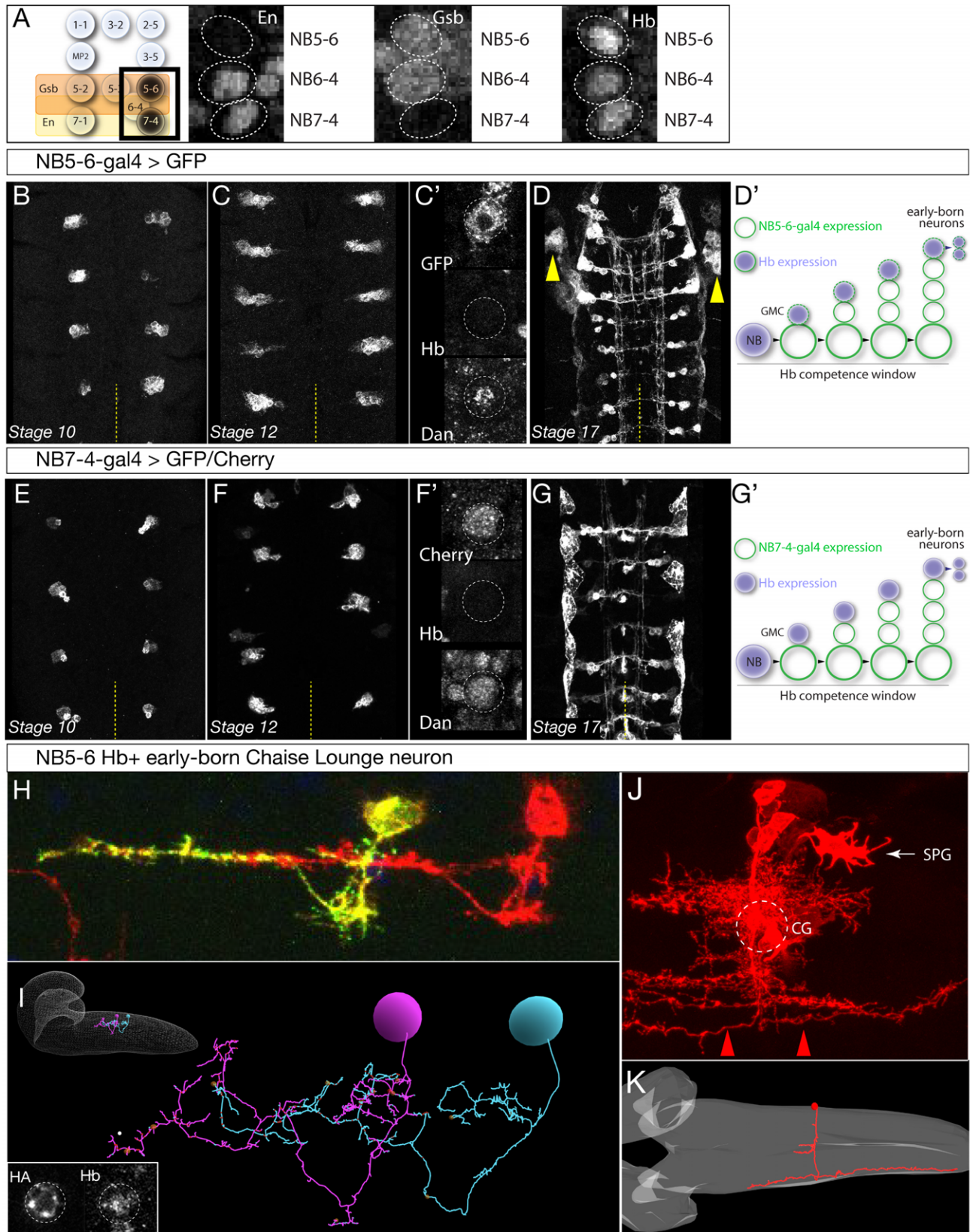
(A) Spatial and temporal patterning. (Left) As neuroblasts delaminate from the neuroectoderm, they experience spatial transcription factors (e.g. Gsb, En, Vnd, Ind, Msh shown) that gives each neuroblast a unique molecular identity. (Middle) TTFs are sequentially expressed in most neuroblasts to specify GMC/neuronal identity based on birth-order. (Right) The integration of spatial and temporal factors specifies lineage-specific neuronal identity.

(B) Independent specification: in this hypothesis, STFs and TTFs bind genomic targets independently, and their combinatorial effect specifies distinct neuroblast identity. In this model, TTF targets are the same in different NBs.

(C) Sequential specification: in this hypothesis, STFs act first to bias or restrict subsequent TTF genomic binding, leading to the production of different neurons from different neuroblasts. In this model, TTF targets are the different in different NBs

(D-F) The TaDa and CaTaDa method. See text for details.



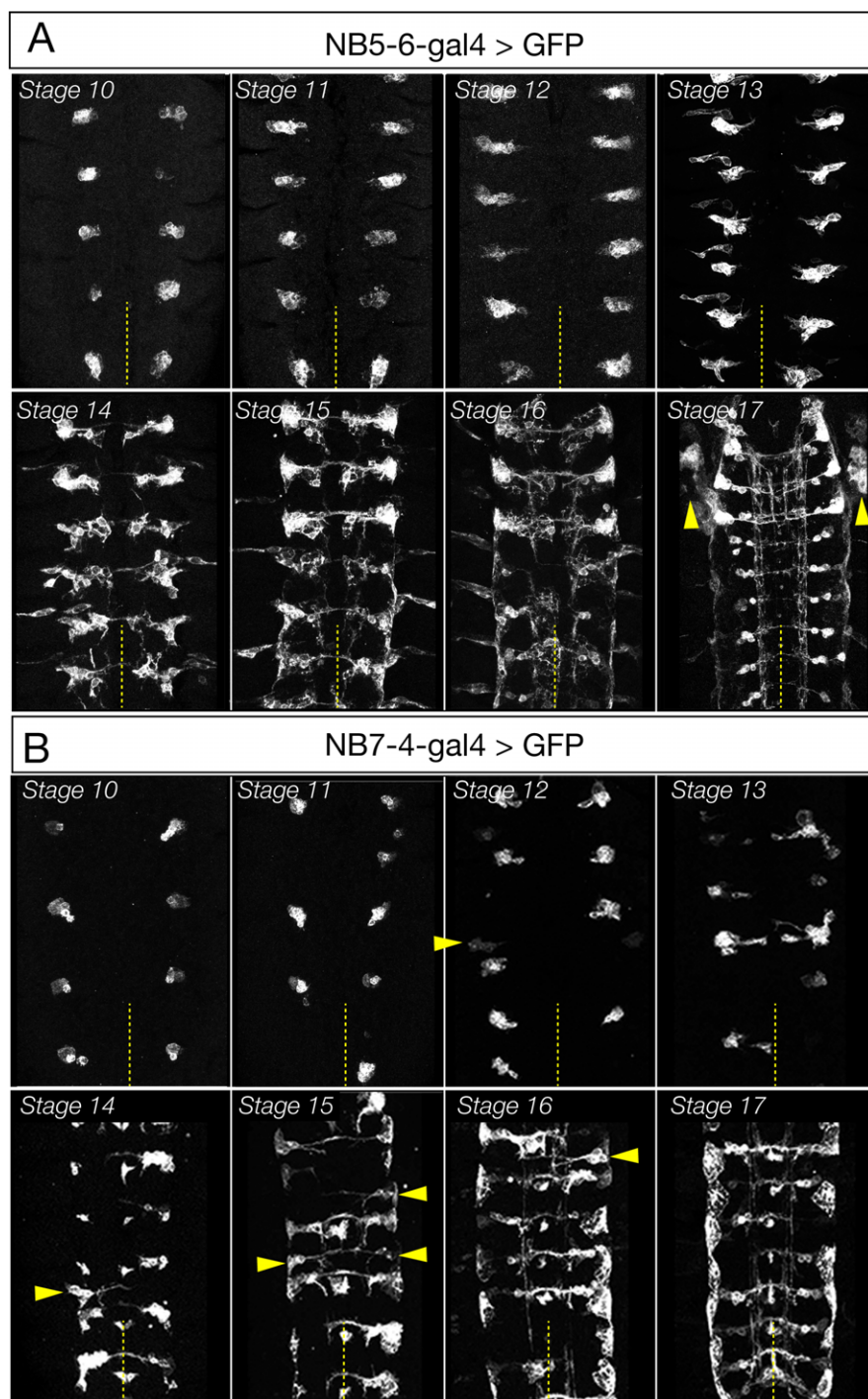


876  
877

**Figure 2. Identification of Gal4 lines specifically expressed in NB5-6 or NB7-4**

878 (A) Left: schematic showing spatial positions of NB5-6 and NB7-4. Right: Immunostaining of  
879 stage 9 embryos showing neuroblast-specific STF expression (En, Gsb) and common TTF  
880 expression (Hb). Genotype: *en-Gal4/UAS-GFP*.  
881 (B-D') *NB5-6-Gal4* is expressed in the NB5-6 lineage from stage 10 until the end of  
882 embryogenesis. Dan is present in NB5-6 through stage 12 (C'). (D') Schematic of NB5-6  
883 expression (green outlines) and Hb expression (purple), see text for details. Note that Gal4  
884 expression is present during the Dan+ Hb competence window. Genotype: *Ibe-K-Gal4/UAS-*  
885 *myr::GFP*.  
886 (E-G') *NB7-4-Gal4* is expressed in the NB7-4 lineage from stage 10 until the end of  
887 embryogenesis. Dan is present in NB5-6 through stage 12 (F'). (G') Schematic of NB7-4  
888 expression (green outlines) and Hb expression (purple), see text for details. Genotype:  
889 *R19B03<sup>AD</sup>/UAS-myrGFP; R18F07<sup>DBD</sup>/+*.  
890 (H-I) NB5-6 early-born Chaise Lounge neurons. Lateral view, anterior, left. (H) Two segmentally  
891 repeated Chaise Lounge neurons labelled by MCFO (*hs-FLP Ibe-K-Gal4 UAS-MCFO*); the  
892 Chaise Lounge neurons are Hb+ (inset). Note the ipsilateral projections. (I) Two segmentally  
893 repeated Chaise Lounge neurons in the EM reconstruction, where they are named A27k. Inset:  
894 outline of CNS with Chaise Lounge neurons shown.  
895 (J-K) NB7-4 early-born G neuron. (J) MARCM clone made with *en-Gal4* labels most or all of the  
896 NB7-4 lineage, including the diagnostic Channel Glia (CG) which are only made by NB7-4  
897 (Schmid et al., 1999; Schmidt et al., 1997). Note the G neuron axon arbors which project the  
898 most laterally in the connective and are both ascending and descending (red arrowheads). SPG,  
899 subperineurial glia. Dorsal view, anterior to left. (J) The G neuron in the EM reconstruction (red).  
900 The neuropil is outlined in gray. Note the lateral axon projection that is ascending and  
901 descending, and the cell body position contacting the neuropil. Also note the two small bilateral  
902 midline processes, which match those of the grasshopper G neuron (Raper et al., 1983).





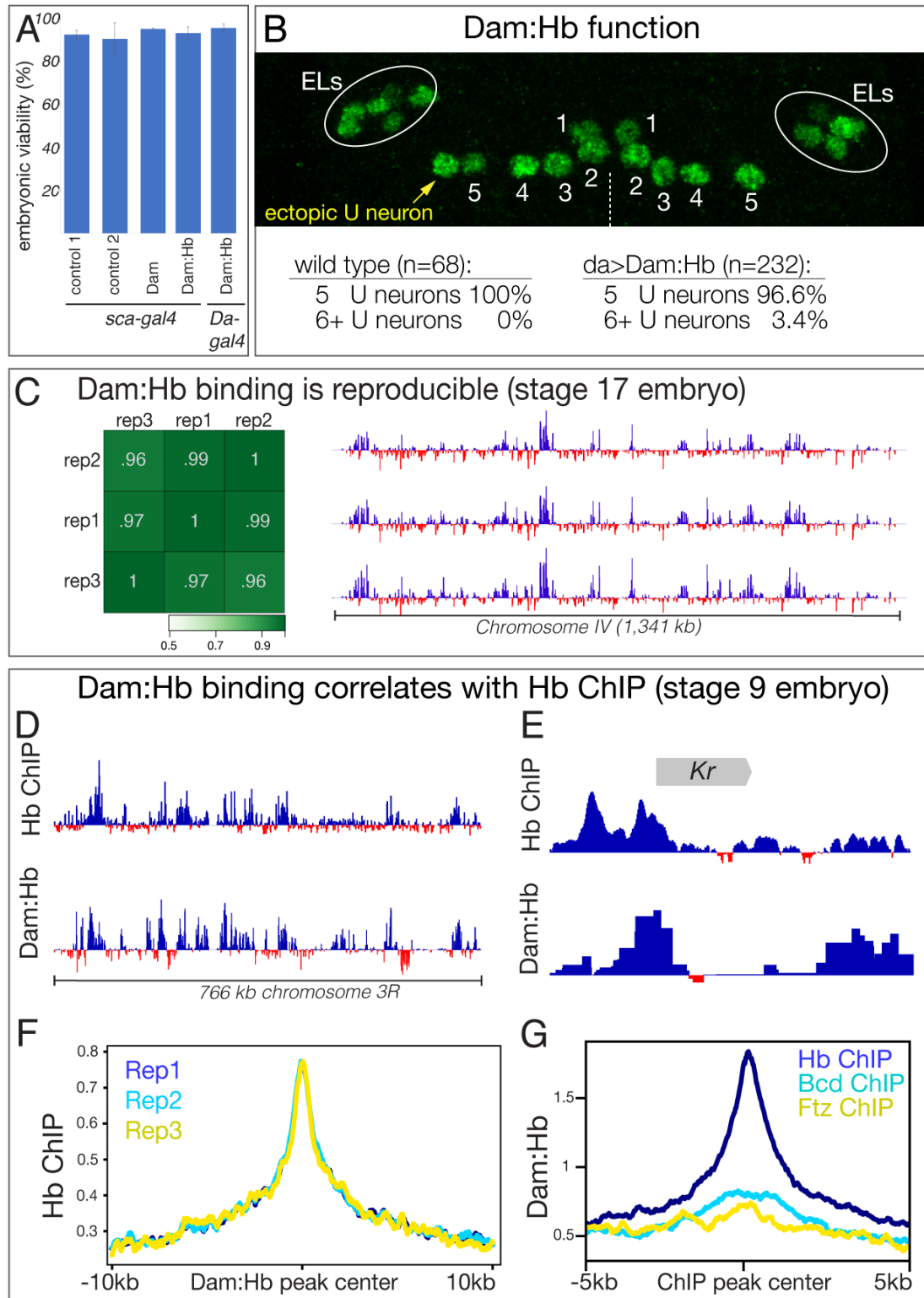
903  
904  
905  
906  
907  
908  
909  
910  
911

**Figure 2 – supplement 1. Expression pattern of NB5-6 and NB7-4 Gal4 lines**

(A) *Lbe-K-Gal4* (*NB5-6-Gal4*) expression is specific to NB5-6 and its progeny from stage 10-stage 16, followed by the addition of salivary gland expression (arrowheads) at stage 17.

Genotype: *Lbe-K-Gal4/UASmyrGFP*

(B) *R19B03<sup>AD</sup>; R18F07<sup>DBD</sup>* split line (*NB7-4-Gal4*) is specific to NB7-4 from stage 10 to the end of embryogenesis. In 6% of hemisegments, the *R19B03<sup>AD</sup>; R18F07<sup>DBD</sup>* line is expressed in the NB5-6 lineage (arrowheads). Genotype: *R19B03<sup>AD</sup>/UAS-myrgFP; R18F07<sup>DBD</sup>/+*.



912  
913  
914  
915  
916  
917

**Figure 3. Generation of a functional, non-toxic Dam:Hb fusion protein**

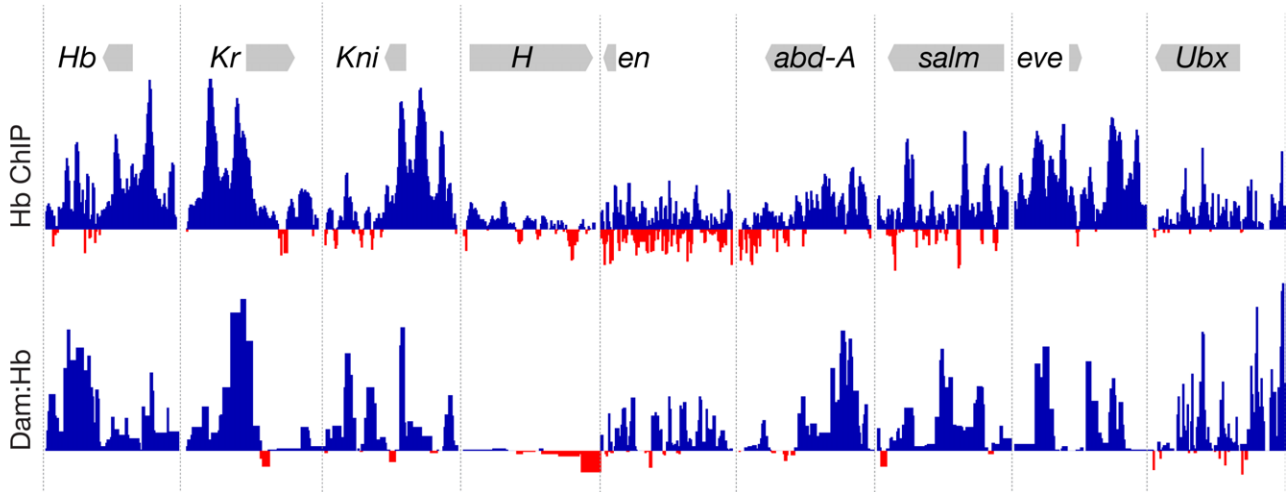
(A) Low level Dam:Hb expression is non-toxic. Control 1, *sca-gal4/sca-gal4*; control 2, *sca-gal4 UAS-HA::UPRT* (Miller et al., 2009); Dam, *sca-gal4 UAS-LT3-Dam*; Dam:Hb, *sca-gal4 UAS-LT3-Dam:Hb*, Dam:Hb, *Da-gal4 UAS-LT3-Dam:Hb* (n=300 for each genotype).

918 (B) Dam:Hb retains Hb function and can induce ectopic Eve<sup>+</sup> U neurons. Anterior up; midline,  
919 dashed line. Left hemisegment shows a single ectopic Eve<sup>+</sup> neuron (yellow) to comprise six total  
920 U neurons, whereas the right hemisegment has the normal five U neurons. Below, quantification.  
921 Wild type (*y w*) represents 68 hemisegments from six embryos; Dam:Hb (*da-Gal4 UAS-LT3-*  
922 *Dam:Hb*, second ORF) represents 8 of 232 hemisegments from 15 embryos with an ectopic U  
923 neuron. ELs, Eve lateral neurons.

924 (C) Dam:Hb binding is reproducible. Left, three biological replicates of genomic binding sites  
925 showing high Pearson correlation coefficients. Right, Dam:Hb binding over 1341 kb on  
926 chromosome IV is highly similar in all three biological replicates. Genotype *da-Gal4 UAS-LT3-*  
927 *Dam:Hb* in stage 17 embryos. Data range: -2.84 – 7.07.

928 (D-G) Dam:Hb-bound loci correlate with Hb ChIP loci. (D) Alignment of Dam:Hb and Hb ChIP  
929 binding sites over 766kb of genomic DNA near the Hb locus, where Hb is known to bind. Data  
930 range for Hb ChIP: -1.01 – 6.23; Data range for Dam:Hb: -2.63 – 5.3. (E) Alignment of Dam:Hb  
931 and Hb ChIP binding sites at the *Krüppel* (*Kr*) locus. Data range for Hb ChIP: -1.66 – 9.04; Data  
932 range for Dam:Hb: -0.63 – 5.68. (F) Dam:Hb peaks for three replicates (blue, cyan, yellow) are  
933 correlated with Hb ChIP signal. Plot shows the Hb ChIP signal +/- 10kb of the center of all the  
934 peaks identified by Dam:Hb analysis in the three replicates. (G) Dam:Hb signal is enriched at  
935 sites of Hb ChIP binding (blue), but not that of Bcd (cyan) or Ftz (yellow). Plot shows the Dam:Hb  
936 signal +/- 5kb of the center of all the peaks identified by ChIP-chip analysis.

937



938

939

940

941

942

943

944

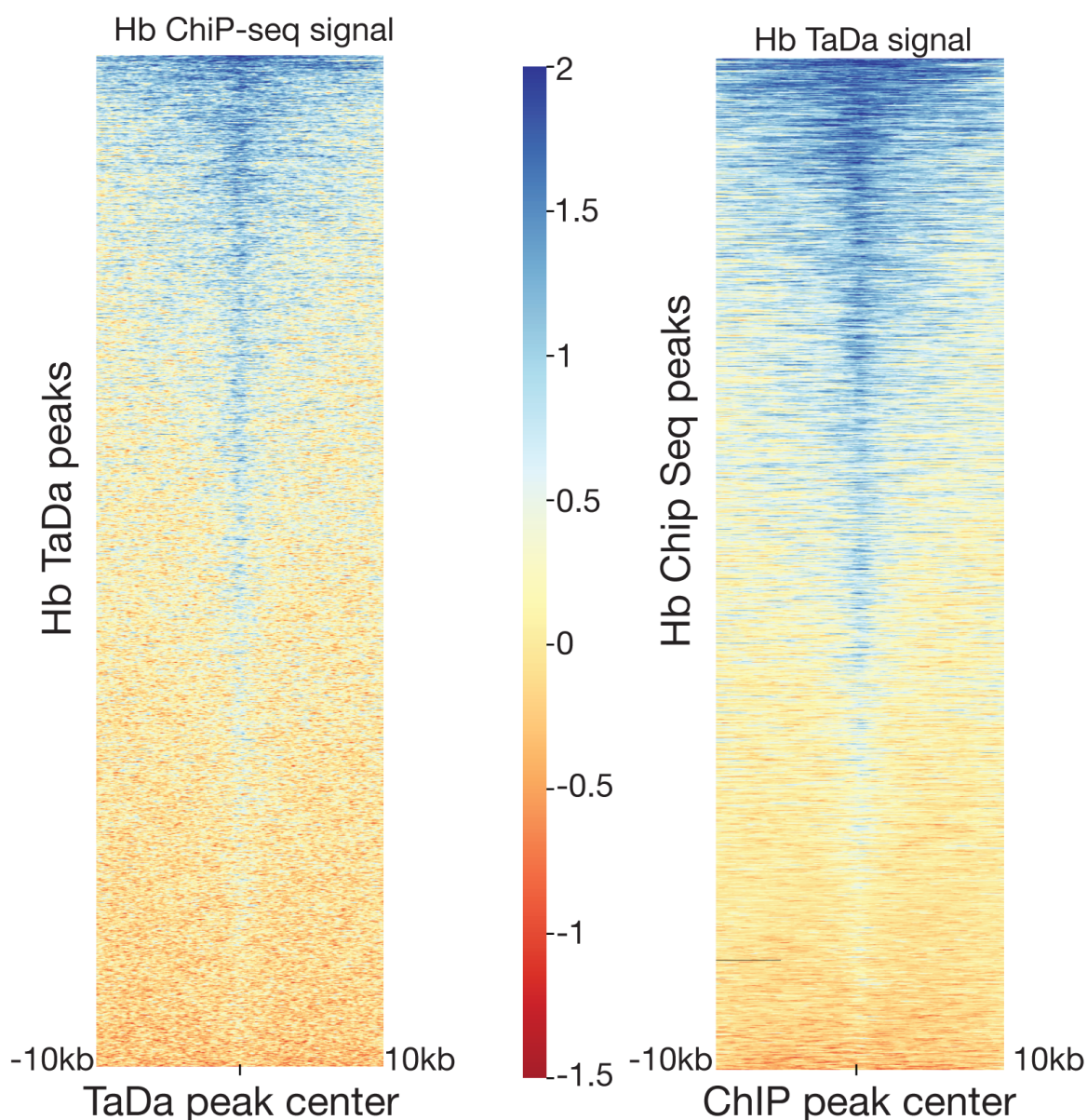
945

946

**Figure 3 – supplement 1. Dam:Hb and Hb-ChIP show similar binding at known Hb target genes**

Top row: Hb ChIP-seq data (Bradley et al., 2010). Bottom row: Dam:Hb data showing the ratio of Dam:Hb binding over Dam alone binding. Dam:Hb enriched binding (blue) and Dam only enriched binding (red). Note the similarity of Hb occupancy as reported by the two techniques at these 9 known targets of Hb (Lyne et al., 2007). Data range for Hb ChIP: -2.45 – 9.04; Data range for Dam:Hb: -0.93 – 6.24.

## Dam:Hb binding correlates with Hb ChIP (stage 9 embryo)








947  
948  
949  
950  
951  
952

### Figure 3 – supplement 2. Dam:Hb and Hb-ChIP binding is correlated.

Left: Increased Hb ChIP signal at Dam:Hb peaks. Right: Increased Dam:Hb signal at Hb ChIP peaks. Heatmap shows the Hb ChIP/Hb TaDa signal +/- 10kb of the center of all the peaks. Each row in the heatmap corresponds an individual peak, and the signal around that peak.



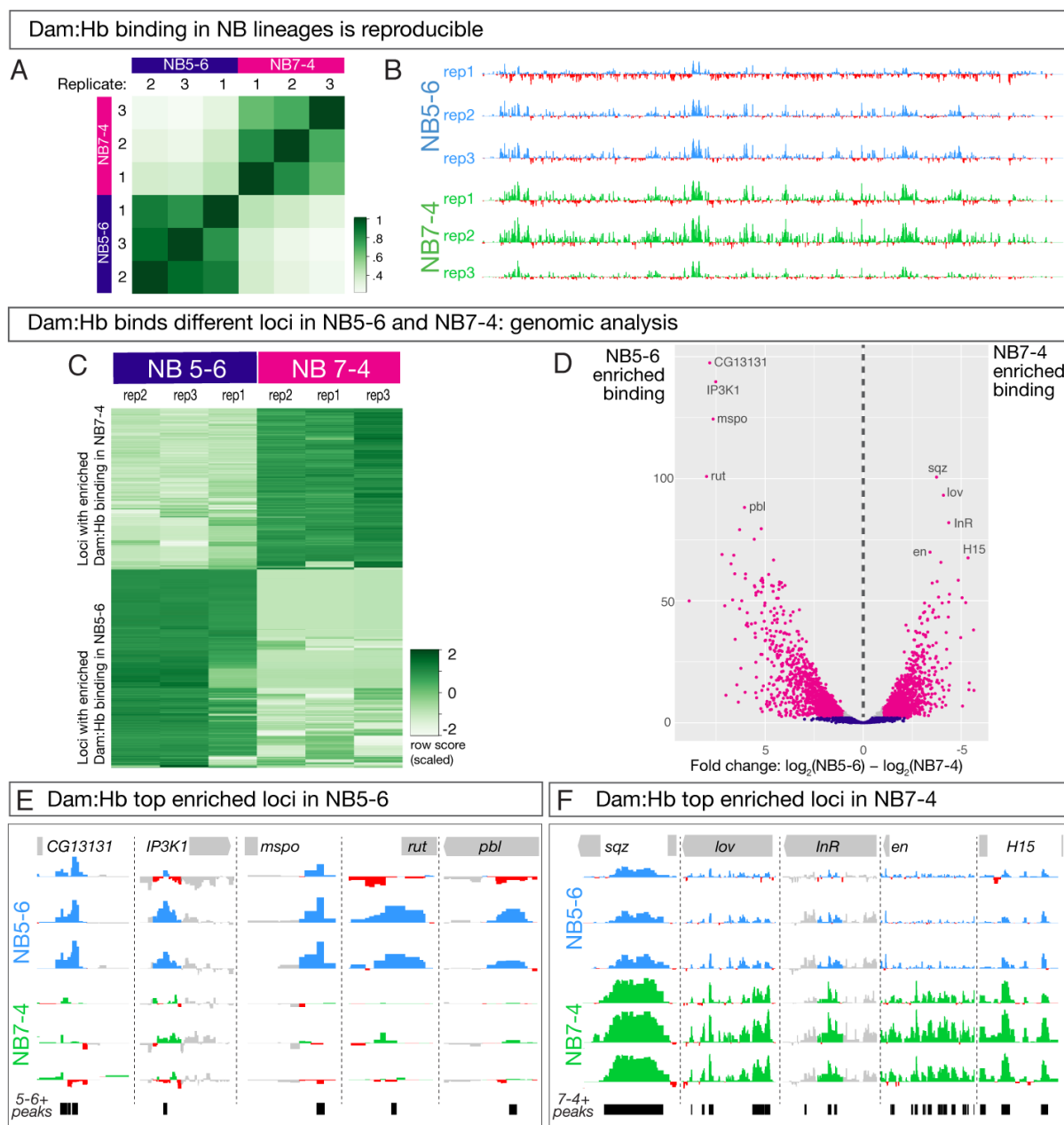
Binding motifs enriched at Dam:Hb peaks						
Rank	Motif	Motif Score	TF	% target	% bkgd	P-value
1		0.93	Hb	29.13	11.95	1e-47
2		0.764	Hb	64.36	48.00	1e-24
3		0.824	Trl	35.44	21.22	1e-24
4		0.713	Six4	30.43	18.52	1e-19
5		0.709	Kr	20.52	10.88	1e-18

953  
954  
955  
956  
957  
958

**Figure 3 – supplement 3. Hb binding motifs are enriched at Dam:Hb bound loci.**

Hb DNA-binding motif (Stanojevic et al., 1989) is enriched in the top 1000 Dam:Hb peaks.

Analysis using the Homer suite of tools (Heinz et al., 2010). Comparable results were obtained when all Dam:Hb peaks were used.



959

960

961

962

963

964

965

966

967

968

969

970

971

**Figure 4. Dam:Hb has distinct genomic binding sites in NB5-6 and NB7-4 lineages**

(A,B) Dam:Hb binding in the NB5-6 lineage and the NB7-4 lineage is reproducible. (A) Three biological replicates of Dam:Hb in each neuroblast lineage are shown, with high Pearson correlation coefficients within each neuroblast replicate, and low correlation coefficients between each neuroblast. (B) Dam:Hb binding over 1,341 kb on chromosome IV is qualitatively similar between lineages. Data range: -3.49 – 8.71.

(C-F) Differential binding data showing Dam:Hb binds different loci in NB5-6 versus NB7-4. (C) A binding affinity heatmap (scaled) showing reads at loci differentially occupied by Dam:Hb in NB5-6 and NB7-4. Loci (rows) are shown for biological replicates of both neuroblasts with greater densities of Dam:Hb binding in darker colours. Note that sites with higher counts in the three NB7-4 replicates (top right) are depleted in the three NB5-6 replicates (top left), and vice versa.

(D) Volcano plot showing differentially occupied loci that are FDR  $\leq 0.01$  in magenta, FDR  $> 0.01$

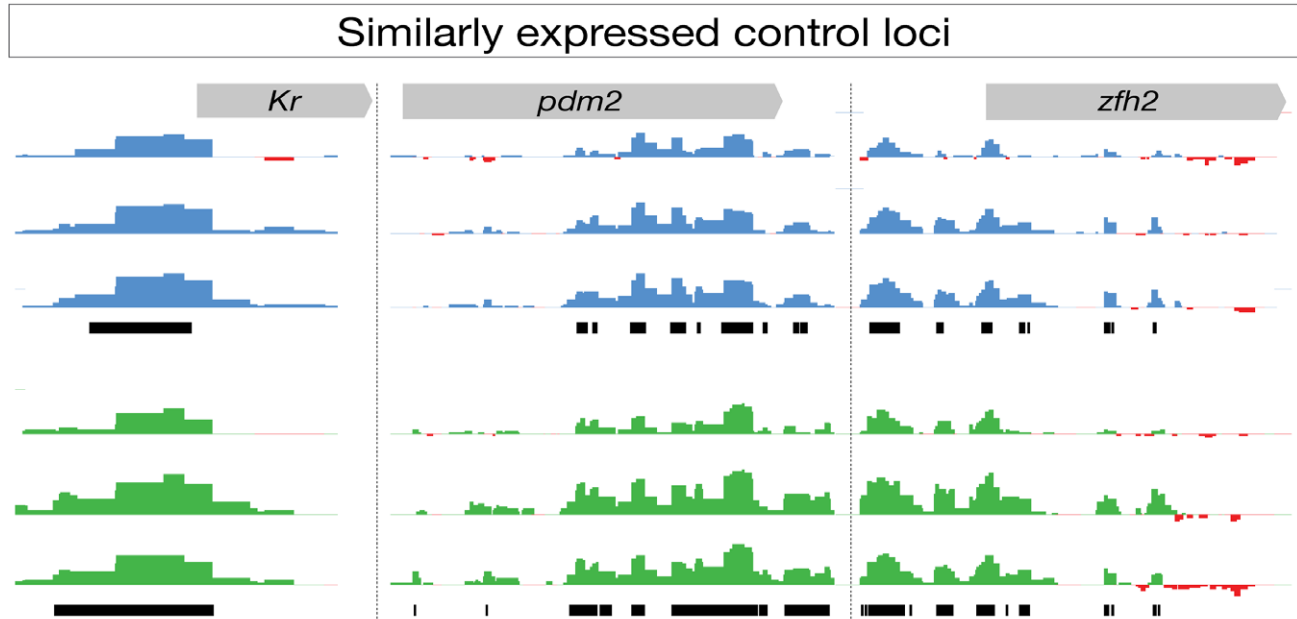
972 in blue, and those that have a fold change of less than 2 in grey. This threshold corresponds to  
973 718 loci in NB5-6 and 504 loci in NB7-4. Genome-wide Hb-bound loci in both neuroblasts were  
974 analysed for differential analysis using DiffBind (Ross-Innes et al., 2012) with DESeq2 and  
975 edgeR and two independent peakcallers with similar results. These plot show DESeq2 results  
976 with the MACS2 peak caller (Zhang et al., 2008).

977 (E,F) The top five enriched Dam:Hb-bound loci are shown for NB5-6 (blue track in F) versus  
978 NB7-4 (green tracks in G) lineages. The black bars represent the loci identified as differentially  
979 bound in the analysis. Data range: -1.9 – 3.96.

980 For all panels, NB5-6 genotype: *NB5-6-Gal4 UAS-LT3-Dam:Hb* or *UAS-LT3-Dam*. NB7-4  
981 genotype: *NB7-4-Gal4 UAS-LT3-Dam:Hb* or *UAS-LT3-Dam*.

982

983



984

985

986

987

988

989

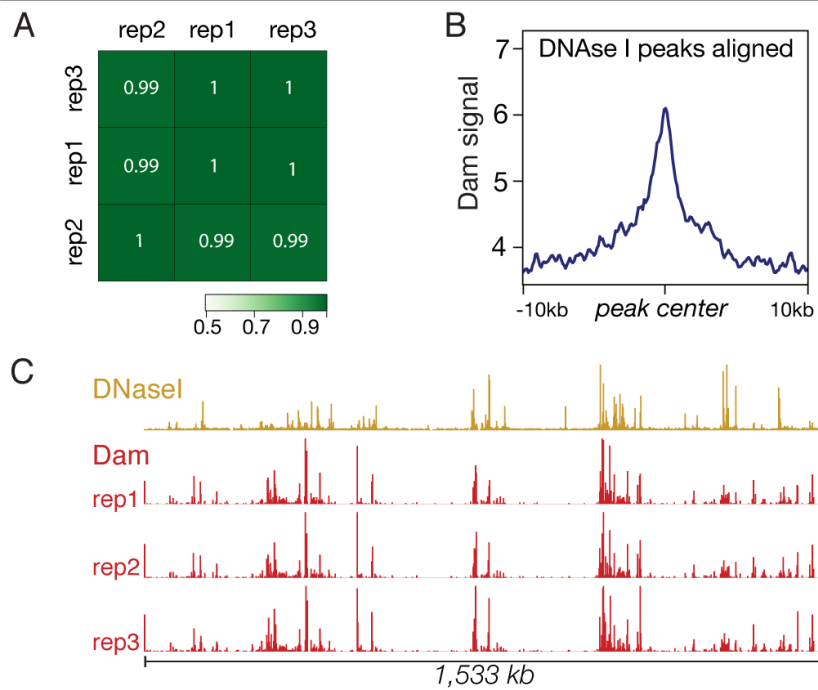
990

991

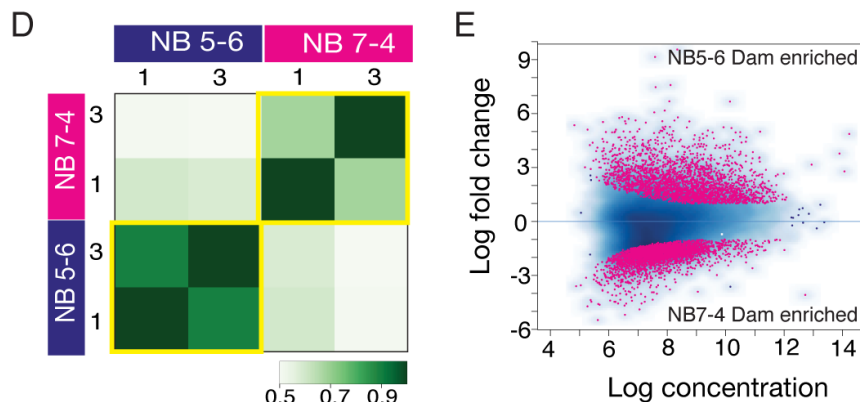
**Figure 4 – supplement 1. Dam:Hb shows similar binding at loci expressed in both NB5-6 and NB7-4**

*Kr*, *pdm2* and *zfh2* are expressed in both NB5-6 and NB7-4 and show similar Dam:Hb binding profiles. Black bars, enriched binding observed in one of the three replicates. Data range: -2.44 – 9.57.

Dam binding is reproducible and correlates with DNase I open chromatin



NB5-6 and NB7-4 lineages have different open chromatin domains



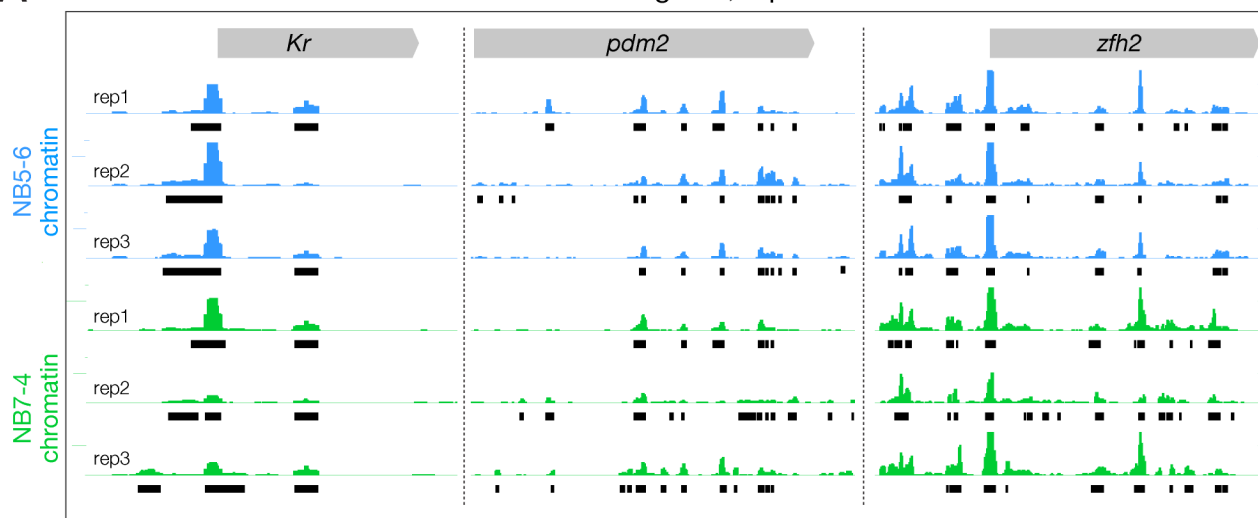
**Figure 5. Dam only binding shows differential open chromatin landscapes in NB5-6 and NB7-4 lineages.**

(A-C) Dam binding is reproducible and correlates with DNase I sites. (A) Three biological replicates are shown, with high Pearson correlation coefficients. (B) Dam binding is enriched at DNase I hypersensitive peaks. (C) Dam binding over 1,533 kb on chromosome 3R is similar in all replicates (red tracks), and similar to DNase I hypersensitivity data (ochre tracks). Data range for Dam: 0 – 50; Data range for DNase I: 0 – 150. Genotype: *Da-Gal4/UAS-LT3-Dam*.

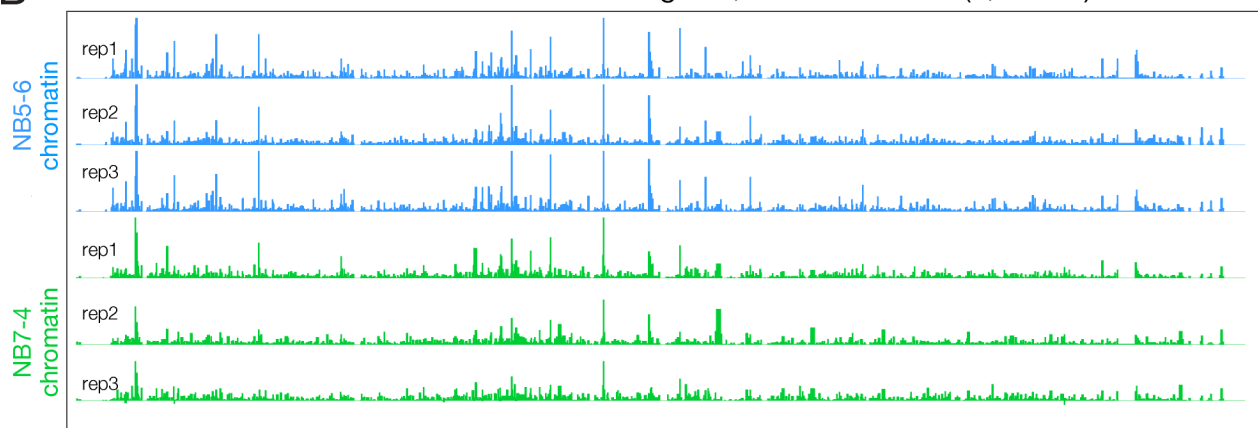
(D-E) Dam binding reveals different open chromatin domains in NB5-6 versus NB7-4. (D) Heat map showing Dam binding sites in NB5-6 have high Pearson correlation coefficients in two replicates, but note the low correlation coefficients between NB5-6 and NB7-4 replicates, showing that each neuroblast has different open chromatin landscapes. (E) Dam binds different loci in the NB5-6 lineage versus the NB7-4 lineage. MA plot showing 3,656 loci enriched for Dam binding in the NB5-6 lineage (top) and 5,084 loci enriched for Dam binding in the NB7-4 lineage (bottom).



**A** Similar chromatin in NB5-6 and NB7-4 at stage 12, representative loci



**B** Similar chromatin in NB5-6 and NB7-4 at stage 12, chromosome IV (1,341 kb)



1007

1008

1009

1010

1011

1012

1013

1014

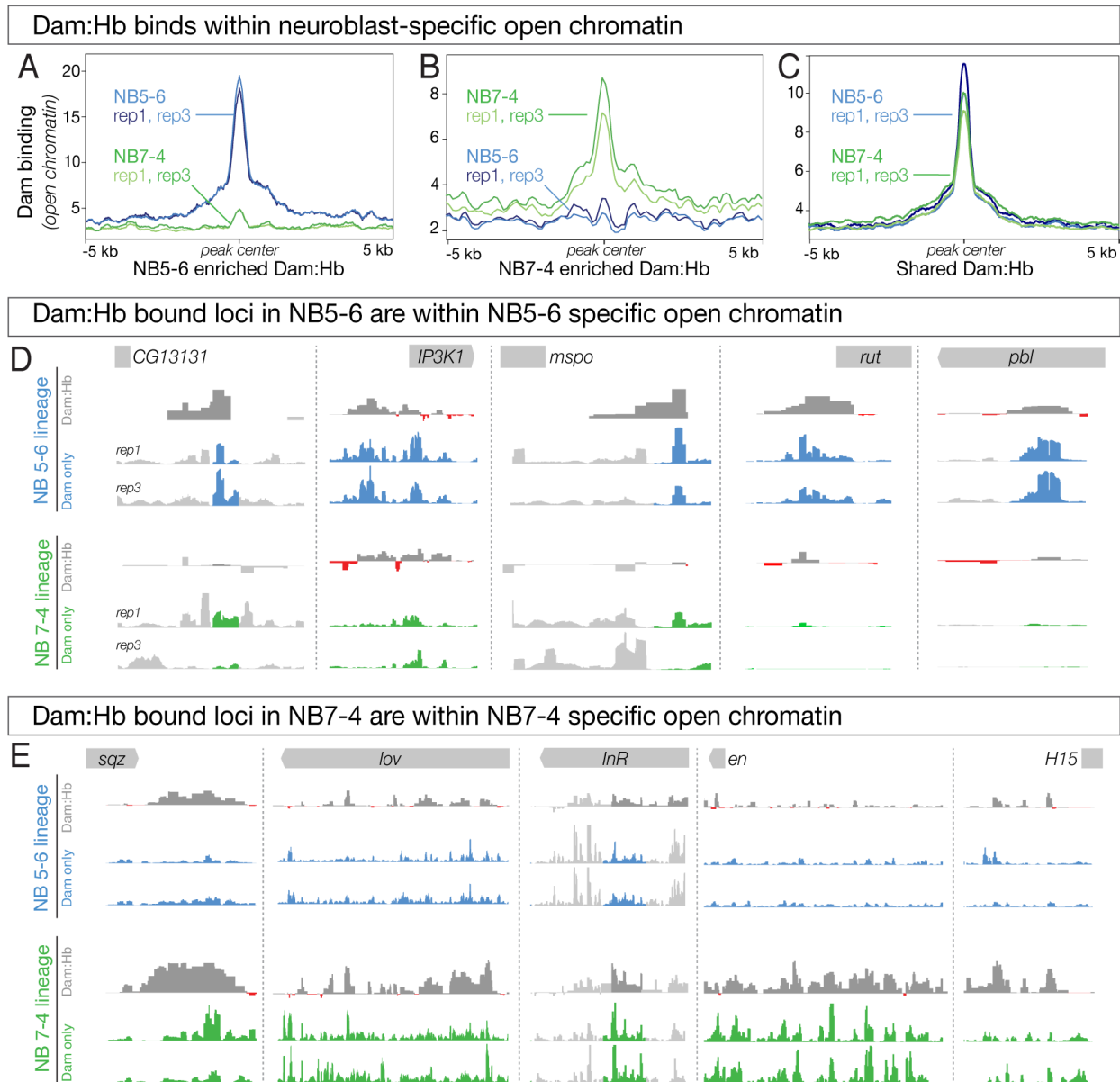
1015

1016

1017

**Figure 5 – supplement 1. Dam only shows similar binding at loci expressed in both NB5-6 and NB7-4**

(A) *Kr*, *pdm2* and *zfh2* are expressed in both NB5-6 and NB7-4 and show similarly open chromatin at these loci. Black bars, enriched binding observed in all three replicates. Data range: 0 - 70. (B) Dam binding over 1,341 kb on chromosome IV in three biological replicates for each neuroblast lineage (NB5-6 lineage – blue tracks; NB7-4 lineage – green tracks), are qualitatively similar within lineages.



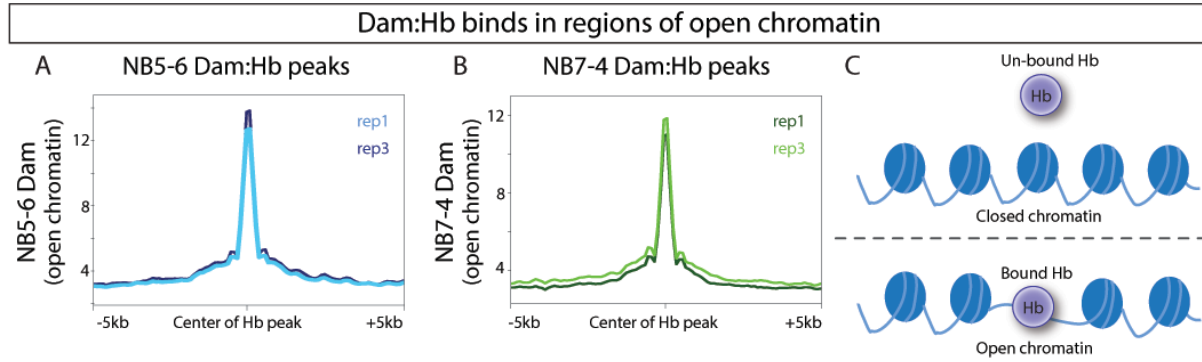
1018  
1019

1020 **Figure 6. Differential chromatin in the 5-6 and 7-4 neuroblast lineages is correlated with**  
1021 **differential Hb occupancy**

1022 (A-C) Dam:Hb binds within neuroblast-specific open chromatin. (A) Dam signal (open chromatin)  
1023 in NB 5-6 (blue lines) and NB 7-4 (green lines) at loci where Dam:Hb binding is enriched in NB5-  
1024 6 over NB7-4. Note that the chromatin is more open in NB5-6 than in NB7-4 at these loci. (B)  
1025 Dam signal (open chromatin) in NB 7-4 (green lines) and NB 5-6 (blue lines) at loci where  
1026 Dam:Hb binding is enriched in NB7-4 over NB5-6. Note that the chromatin is more open in NB7-  
1027 4 than in NB5-6 at these loci. (C) Dam signal (open chromatin) at loci similarly occupied by Hb in  
1028 both NB5-6 and NB7-4 lineages.

1029 (D) The top five Dam:Hb enriched loci in NB5-6 are in regions of NB5-6 open chromatin (blue  
1030 tracks); however, in NB7-4 these loci are not in open chromatin (Dam; green tracks), and are not  
1031 bound by Dam:Hb. Rows from top to bottom: genomic locus, Dam:Hb enrichment in NB5-6, Dam  
1032 only enrichment in two replicates in NB5-6, Dam:Hb enrichment in NB7-4, and Dam only

1033 enrichment in two replicates in NB7-4. Data range for *IP3K1*, *rut*, *pbl* is 0-109; data range for  
1034 *CG13131* and *mspo* is 0-15.  
1035 (E) The top five Dam:Hb enriched loci in NB7-4 are in regions of open chromatin in NB7-4 (green  
1036 tracks); however, in NB5-6 these loci are not in open chromatin (Dam; blue tracks) and are not  
1037 bound by Dam:Hb. Rows from top to bottom: genomic locus, Dam:Hb enrichment in NB5-6, Dam  
1038 only enrichment in two replicates in NB5-6, Dam:Hb enrichment in NB7-4, and Dam only  
1039 enrichment in two replicates in NB7-4. Data range for *sqz*, *InR* and *en* is 0-35; data range for *lov*  
1040 and *H15* is 0-20.  
1041



1042

1043

1044

1045

1046

1047

1048

1049

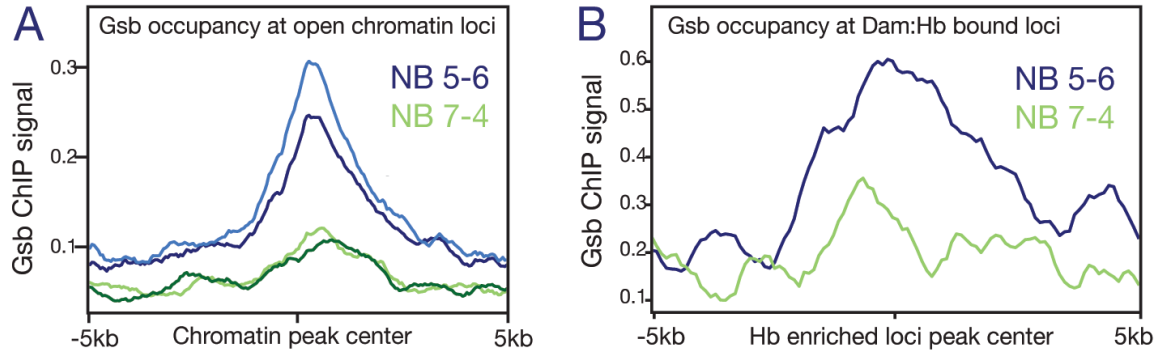
1050

1051

1052

**Figure 6 – supplement 1. Dam:Hb binding is biased towards regions of open chromatin**

Dam binding (open chromatin) is correlated with Dam:Hb binding in each neuroblast lineage. (A) Dam signal (open chromatin) of the two replicates of NB5-6 is plotted at the region of Dam:Hb occupancy in the NB5-6 lineage. (B) Dam signal (open chromatin) of the two replicates of NB7-4 is plotted at the region of Dam:Hb occupancy in the NB7-4 lineage. Note the correspondence of open chromatin (increased Dam signal) at the regions of Hb occupancy in both NB lineages. These results are consistent with preferential Hb binding at regions of open chromatin. (C) Schematic showing preferential binding of Dam:Hb at sites of open chromatin.



1053

1054

1055

1056

**Figure 7. Gsb binding is enriched at open chromatin and Dam:Hb bound loci in NB5-6, but not NB7-4.**

1057 (A) Gsb ChIP-chip signal at the regions of Dam-bound (open) chromatin; note the enrichment in

1058 NB5-6 (blue lines) but not NB7-4 (green lines).

1059 (B) Gsb ChIP-chip signal at the regions of Dam:Hb bound loci; note the enrichment in NB5-6

1060 (blue lines) compared to NB7-4 (green lines).

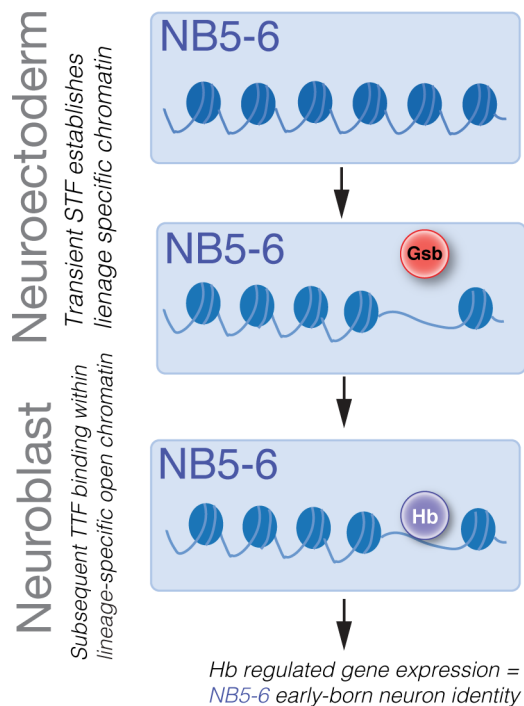
1061

1062

1063

1064

1065



1066

1067

1068

1069

1070

1071

1072

**Figure 8. Sequential specification integrates spatial and temporal cues to generate diversity in Drosophila embryonic NB lineages.**

Transient expression of spatial factors in the neuroectoderm (e.g. Gsb in row 5) establishes lineage-specific chromatin landscapes (e.g. NB5-6 lineage). Subsequently, TTFs (e.g. Hb) in the NB can access different genomic targets to regulate different genes in spatially distinct NB lineages. This results in the specification of different neural fates in different NB lineages.