nature genetics



# RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo

Julia Zeitlinger[1,7], Alexander Stark[2,3], Manolis Kellis[2,3], Joung-Woo Hong[4], Sergei Nechaev[5], Karen Adelman[5], Michael Levine[4] & Richard A Young[1,6]

**It is widely assumed that the key rate-limiting step in gene activation is the recruitment of RNA polymerase II (Pol II) to the core promoter[1]. Although there are well-documented examples in which Pol II is recruited to a gene but stalls[2–12], a general role for Pol II stalling in development has not been established. We have carried out comprehensive Pol II chromatin immunoprecipitation microarray (ChIP-chip) assays in *Drosophila* embryos and identified three distinct Pol II binding behaviors: active (uniform binding across the entire transcription unit), no binding, and stalled (binding at the transcription start site). The notable feature of the ~10% genes that are stalled is that they are highly enriched for developmental control genes, which are either repressed or poised for activation during later stages of embryogenesis. We propose that Pol II stalling facilitates rapid temporal and spatial changes in gene activity during development.**

Pol II stalling is probably best studied at heat shock genes in *Drosophila*, where Pol II engages in transcription but pauses immediately downstream of the transcriptional start site[2,3,13]. Upon activation by heat shock, Pol II is able to rapidly transcribe these genes. Regulation of Pol II activity after recruitment has also been described in bacteria[14], yeast[11] and mammalian cell lines[2,5–10], and it includes instances where Pol II is found in an inactive preinitiation complex[15,16]. We will collectively refer to inactive Pol II near the transcription start site as stalled Pol II.

To determine at which genes Pol II stalling occurs during development, we analyzed global Pol II occupancy in whole *Drosophila* embryos. Although this is one of the few systems in which genomics approaches can easily be applied to developmental questions, interpretation is complicated by the occurrence of multiple tissues. To reduce the complexity, we used *Toll*$^{10b}$ embryos (2–4 h after fertilization), a well-characterized mutant that contains a homogeneous population of mesodermal precursor cells at the expense of neuronal and ectodermal cells[17–23]. In *Toll*$^{10b}$ mutants, mesodermal genes are uniformly activated, whereas genes required for the development of ectodermal and neural tissues are repressed throughout the embryo[17–19]. Previous whole-genome microarray experiments have identified the transcript levels of all genes in these mutants[21,22]. To distinguish between stalled and active Pol II, we used a mixture of antibodies that recognizes both the initiating and elongating forms of Pol II (see Methods) and carried out whole-genome ChIP-chip assays as previously described[23].

The results show that many genes known to be repressed in *Toll*$^{10b}$ embryos show notably high Pol II signal near the transcription start site (**Fig. 1a–d**). In some cases, the prominent Pol II peak was tightly restricted to the promoter region (for example, at the tail-up (*tup*) gene; **Fig. 1a**), whereas at other genes Pol II was also found at low abundance throughout the transcription unit (for example, the *sog* and *brk* genes; **Fig. 1c,d**). This is consistent with previous evidence that some genes, such as *sog*, are transiently activated but then repressed at later stages[24], whereas others, such as *tup*, are never activated in *Toll*$^{10b}$ mutants[21,22].

The Pol II profiles of repressed genes are clearly distinct from those of active genes (**Fig. 1e,f**). For example, the *stumps* (also known as *Hbr*) gene, which encodes a fibroblast growth factor (FGF) receptor specifically expressed in mesodermal precursors (**Fig. 1e**), and ribosomal genes such as *RpL3* (**Fig. 1f**) show uniformly high levels of Pol II throughout the transcription unit. Furthermore, genes that are silent in the early embryo simply lack Pol II binding altogether (**Fig. 1g,h**). Thus, there appear to be three distinct classes of genes: those with Pol II distributed throughout the transcription unit, those with preferential enrichment of Pol II at the transcription site and those that lack Pol II binding altogether.

To further characterize these three groups, we developed a principled method that classifies genes on the basis of their Pol II enrichment profiles (**Fig. 2** and **Supplementary Note** online). In an

[1]Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts 02141, USA. [3]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [4]Department of Molecular Cell Biology, Center for Integrative Genomics, University of California, Berkeley, California 94720, USA. [5]Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709, USA. [6]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [7]Present address: Stowers Institute for Medical Research, 1000 East 50th St., Kansas City, Missouri 64110, USA. Correspondence should be addressed to M.L. (mlevine@berkeley.edu) or R.A.Y. (young@wi.mit.edu).

Received 10 May; accepted 7 September; published online 11 November 2007; doi:10.1038/ng.2007.26

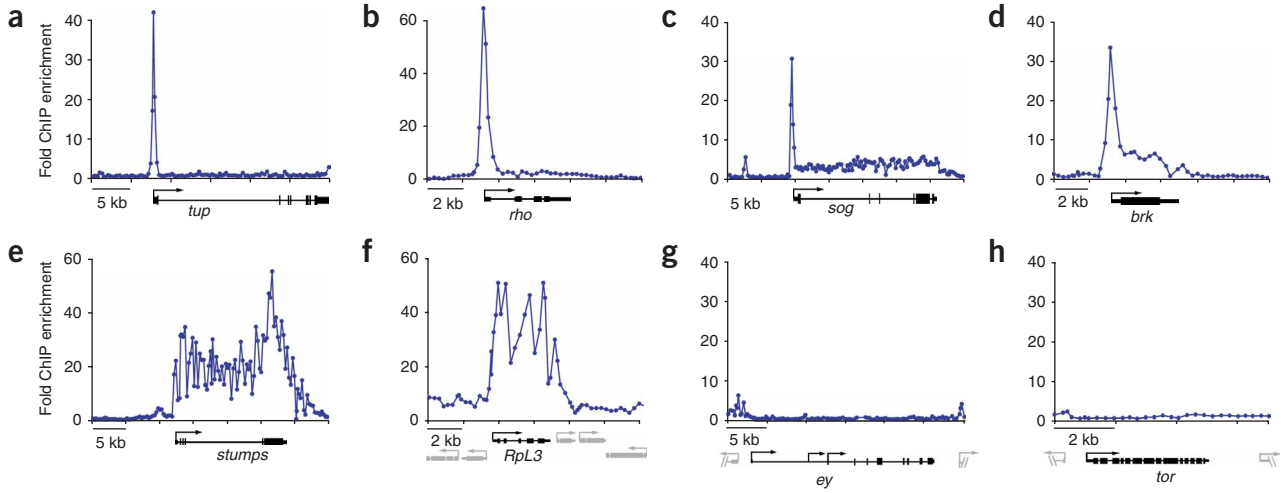


**Figure 1** Different classes of Pol II binding profiles. ChIP-chip assays were carried out with 2–4 h *Toll*$^{10b}$ embryos using antibodies that recognize both the initiating and the elongating forms of Pol II. The enrichment ratios of Pol II are shown on the *y* axis. (**a**–**d**) Binding patterns across genes that are repressed in *Toll*$^{10b}$ embryos. All four genes show high Pol II signals near the transcription start sites. At some genes, such as *tup* (**a**), Pol II is tightly restricted to this region, whereas at other genes, including *sog* (**c**) and *brk* (**d**), Pol II is also detected at lower signals throughout the transcription unit. (**e**,**f**) Pol II is uniformly distributed across the transcription units of genes that are actively transcribed. The *stumps* (also called *Hbr* or *Dof*—downstream of FGF) gene (**e**) is specifically activated in mesodermal precursor cells, whereas *RpL3* (**f**) is a highly expressed ribosomal gene. (**g**,**h**) No Pol II binding is found at many genes that are inactive during embryogenesis. The eyeless (*ey*) gene (**g**) is expressed during eye development at larval stages but not in the early embryo. Likewise, the torso (*tor*) gene (**h**) is active only during oogenesis but not in the early embryo.

analysis similar to one carried out in *Escherichia coli*[14], we calculated the ratio between Pol II enrichment at the transcription start site versus internal regions of the transcription unit (**Fig. 2**). We were able to assign 76% of the protein coding genes (10,220 of 13,448 genes) to one of the three classes (**Fig. 2b**). At least 27% of all genes had an active Pol II profile in which Pol II was detected uniformly throughout the transcription unit. At least 12% of all genes (1,614 of 13,448) showed disproportionate accumulation of Pol II near the transcription start site. Among this group, Pol II was tightly restricted to the transcription start site at 62% of genes. At the remaining 38% of these genes, Pol II was also detected within the transcription unit, presumably because these genes—such as *sog*—are expressed at low levels in at least a subset of cells during the time frame of the analysis (2–4 h after fertilization). Finally, 37% of all genes lacked Pol II binding altogether.

Several lines of evidence confirm that the ~1,600 genes with disproportionate enrichment of Pol II at the transcription start site have a form of stalled Pol II (**Fig. 3**). First, all heat shock genes, which provide the classical example of Pol II stalling[3,13], fall into this class (**Fig. 3a**). Second, the Pol II peaks map an average of ~50 bp downstream of the transcription start site, consistent with the location of stalled Pol II at heat shock genes[3,4,13] (**Fig. 3b**). Because this is an average profile, it is possible that a fraction of Pol II occupancy comes from inactive preinitiation complexes. However, the majority of detected Pol II signal seems to come from Pol II that is stalled downstream of the transcription start site. Third, Pol II stalling at these genes is consistent with comprehensive expression analysis using whole-genome tiling arrays[22]. Genes with Pol II tightly restricted to the transcription start site are either silent or only weakly expressed in *Toll*$^{10b}$ mutants (**Fig. 3c**). In contrast, genes with similar levels of Pol II binding but uniform distribution throughout the transcription unit are expressed at substantial levels in these mutants (**Fig. 3c**). Finally, we used permanganate footprint assays as an independent method to confirm stalled Pol II at selected genes[4,12]. For example, the *rho* gene showed clear permanganate sensitivity downstream of the transcription start site (+37 bp), consistent with the Pol II stalling profile seen in *Toll*$^{10b}$ mutants (**Fig. 3d** and **Fig. 1b**).

There are considerable differences in the expression and functions of genes in the active, stalled or no Pol II classes based on *in situ* expression patterns (ImaGO database)[25] and functional annotations (Gene Ontology database)[26] (**Fig. 4**). The set of genes with stalled Pol II is highly enriched for developmentally regulated genes, particularly those expressed in ectodermal and neuronal precursor cells (**Fig. 4a**).

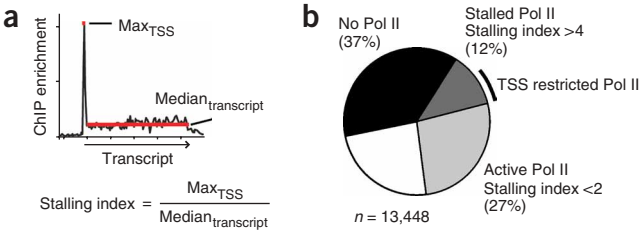


**Figure 2** Whole-genome analysis of Pol II binding. (**a**) Genes were assigned to one of three classes—stalled Pol II, active Pol II or no Pol II—on the basis of their stalling index. The stalling index is the ratio between the maximum enrichment near the transcription start site ($Max_{TSS}$; ±300 bp) and the median enrichment of the probes distributed across the transcription unit ($Median_{transcript}$; excluding the first 600 bp). Stalling index values of >4 qualified as 'stalled Pol II', whereas stalling index values <2 qualified as 'uniform (active) Pol II'. If no probe within the TSS region was significantly bound, the gene was assigned to the 'no Pol II' category. (**b**) Over 76% of all protein-coding genes could be assigned to one of three categories on the basis of stalling index values: 12% have stalled Pol II, 27% show the active form of Pol II and 37% have no Pol II. Among the genes with stalled Pol II, 62% have Pol II tightly restricted to the transcription start site.

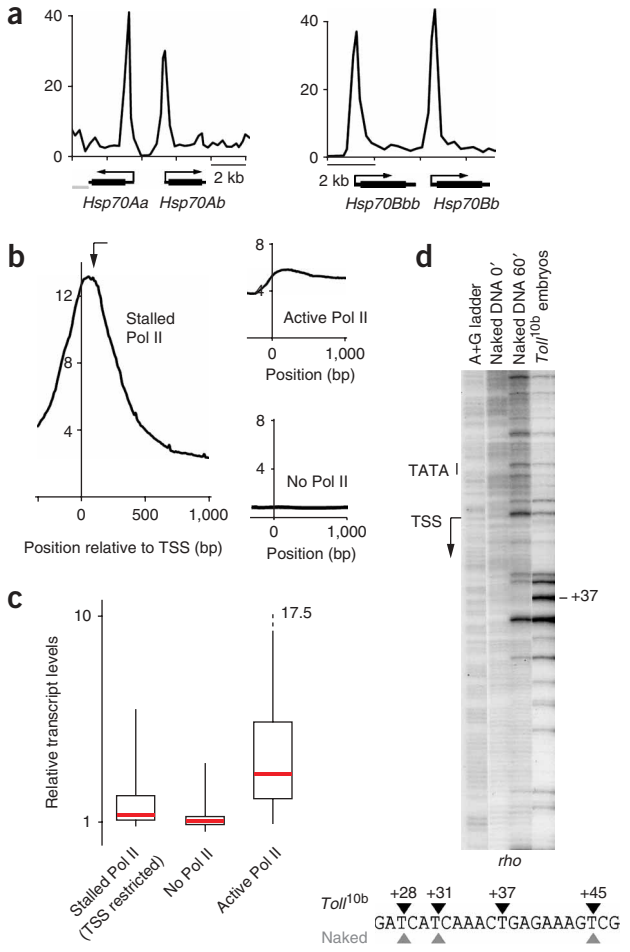


**Figure 3** Confirmation of the class of genes with Pol II stalling. (**a**) The group of stalled genes contains all known heat shock genes where Pol II stalling has been well documented. (**b**) Metagene analysis shows that the average peak of 'stalled Pol' II is ~50 bp downstream of transcription (arrow). The profile for 'no Pol II' and 'active Pol II' is shown as a comparison. (**c**) Analysis of the transcript levels confirms that genes with the stalled Pol II profile (tightly restricted to the transcription site) are either silent or expressed at low levels. Genes that show Pol II enrichment at comparable levels throughout the transcription unit (active Pol II) are expressed at significantly higher levels. Genes with no Pol II are not expressed. The transcript levels are represented as box and whiskers plot of the fold ratios (measured by whole-genome tiling arrays[22]). The box represents the 25th and 75th percentiles, with the median as a red bar. The whiskers refer to the 1st and 99th percentiles. The scale on the *y* axis is a log scale. (**d**) A permanganate footprint assay of the *rho* gene confirms stalled Pol II downstream of the transcription start site. Genomic sequences of A+G are shown as marker (lane 1). In comparison with purified genomic DNA, which was either not treated (lane 2) or treated (lane 3) with $KMnO_4$, a prominent hypersensitive T residue is detected in *Toll*$^{10b}$ mutant embryos (lane 4), implying the existence of a transcription bubble at the region around +37 *in vivo*. The bottom panel shows actual sequences from +26 to +47 of the *rho* locus (relative to TSS as +1 bp).

Consistent with these results, genes with stalled Pol II are highly enriched for functions in development, including neurogenesis, ectoderm development and muscle differentiation (**Fig. 4b**). Many of these genes encode sequence-specific transcription factors (Hox, T-box, bHLH, zinc fingers and HMG) and components of cell signaling pathways (FGF, Wnt, Notch, EGF, TGFβ, JNK and TNF) (**Supplementary Note**).

In contrast, the set of genes with uniform Pol II binding is highly enriched for ubiquitously expressed genes (**Fig. 4a**), which function mostly in metabolism and cell proliferation (**Fig. 4b**). The set of genes that lacks Pol II binding is highly enriched in genes that show no staining in whole-embryo *in situ* hybridizations, confirming that they are not expressed during early embryogenesis (**Fig. 4a**). Many of these genes encode proteins that have functions in adult cells, such as cuticle proteins or proteins required for vision (**Fig. 4b**).

Pol II stalling could reflect two nonexclusive developmental functions. It could be indicative of active transcriptional repression, or it could prepare genes for activation at later stages of embryogenesis. The second model is particularly attractive, because Pol II stalling has already been shown to prepare heat shock genes for rapid induction[3,13]. We found evidence for both models.

Pol II stalling is particularly prevalent among genes expressed in the neuroectoderm and dorsal ectoderm, which are repressed in *Toll*$^{10b}$ embryos. To test whether Pol II stalling is specific for repressed genes, we examined the Pol II profile of these genes in mutant embryos

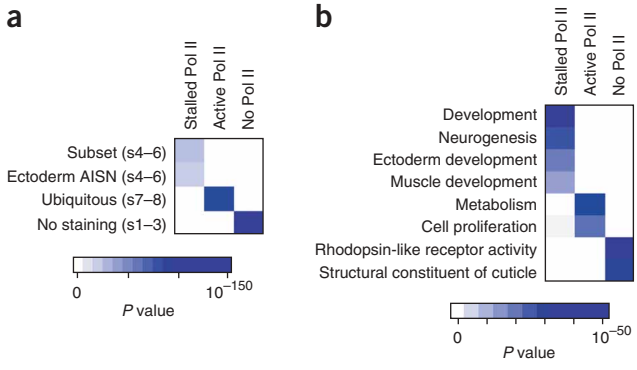


**Figure 4** Functional analysis of the three classes of genes. (**a**,**b**) Representative categories of enrichment among genes with stalled Pol II, active Pol II or no Pol II for gene sets are shown for the ImaGO database (**a**), which contains the *in situ* expression patterns of a substantial fraction of all protein coding genes in the *Drosophila* genome, and the gene sets from the biological process categories in the Gene Ontology (GO) database (**b**). The scale bars below indicate the significance of each test (hypergeometric distribution). Genes containing stalled Pol II are significantly enriched for genes expressed in a subset of cells and in those of developing ectoderm (AISN = *anlage in statu nascendi*) at the time of the analysis (120–240 min). The stage of the enriched categories is indicated in parenthesis (s1–3 = ~0–100 min, s4–6 = ~100–200 min, s7–8 = ~200–250 min). The genes expressed as subset (s4–6) are largely identical to those of the ectoderm AISN (s4–6) category, because mesoderm, neurectoderm and dorsal ectoderm are specified at that stage. Functional analysis confirms that genes with stalled Pol II are enriched for genes involved in development, in particular those required for neurogenesis, ectoderm differentiation and muscle development. Genes showing active Pol II are enriched for genes that are ubiquitously expressed in developing embryos. They are enriched for functions that mediate cell proliferation and metabolic functions such as protein and nucleotide metabolism. Genes lacking Pol II tend to be inactive during embryogenesis, and they are deployed at later stages of the life cycle, such as cuticle function and vision.

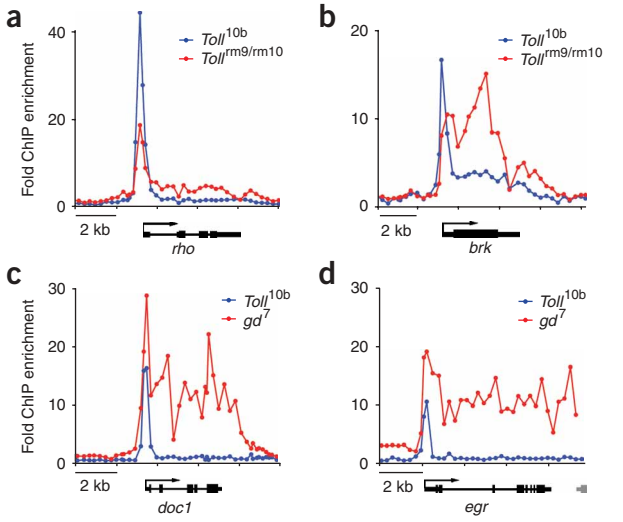


**Figure 5** Pol II profile comparison of genes in the repressed versus active state. (**a**–**d**) Pol II ChIP enrichment ratios are shown for genes that are repressed in *Toll*$^{10b}$ embryos but are active in either *Toll*$^{rm9/rm10}$ embryos (**a**,**b**) or *gd*$^{7}$ embryos (**c**,**d**) at 2–4 h after fertilization. The enrichment ratios of Pol II are shown on the *y* axis: active state (red) and repressed state (blue). The results show that the degree of Pol II stalling is dependent on the gene's activity, with some genes showing a complete switch between the stalled and active forms of Pol II.

in which they are active. For this, we used two well-defined mutants, *Toll*$^{rm9/rm10}$ and *gd*$^{7}$ (2–4 h), in which cells adopt neurecto-dermal and dorsal ectodermal fates, respectively[17,21,22,27]. Indeed, at these genes, Pol II is redistributed into the transcription unit in these mutants (**Fig. 5**), and some genes now show the active Pol II profile (**Fig. 5b–d**). These results indicate that Pol II stalling is associated with cell-type specific repression and is subject to dynamic changes during development.

Previous studies have shown that the repression of a large set of genes in *Toll*$^{10b}$ embryos depends on snail, a well-studied repressor that is constitutively expressed in *Toll*$^{10b}$ embryos but not in *Toll*$^{rm9/rm10}$ and *gd*$^{7}$ embryos[18,19,21–23,28]. We found a statistically significant association between repression by snail and Pol II stalling. For example, among the 139 genes that are occupied by snail[23] and show reduced expression in the *Toll*$^{10b}$ mutant[21], 54% have stalled Pol II, whereas only 19% of all genes with reduced expression show Pol II stalling ($P < 10^{-23}$, **Supplementary Note**). This suggests that Pol II stalling in *Toll*$^{10b}$ embryos may be regulated by snail. A role of developmental repressors in regulating Pol II stalling is also consistent with a recent study of *Drosophila* segmentation[12].

Multiple lines of evidence suggest that Pol II stalling also occurs at genes that are poised for activation in older embryos. Genes with stalled Pol II are highly over-represented among genes that are rapidly induced within 12 h after the time frame of our analysis ($P < 10^{-27}$, **Supplementary Note**). Moreover, genes with stalled Pol II are enriched for genes expressed in the derivatives of the mesoderm precursors present in *Toll*$^{10b}$ mutants, such as the developing heart and muscle cells ($P < 10^{-15}$, **Supplementary Note**). These genes, such as Drop (*Dr*) and *bap*, are not yet activated at the time frame of the analysis[29,30], but they nonetheless show high levels of Pol II near the transcription start site (**Supplementary Note**).

To confirm that muscle genes indeed show stalled Pol II before activation, we carried out permanganate assays on wild-type *Drosophila* embryos at 2–4 h after fertilization (**Fig. 6**). *Dr* and *lbe* showed a clear permanganate footprint downstream of transcription. These footprints were specific to the early embryo stage, as S2 cells, a cell line derived from older embryos, did not show a permanganate footprint under the same conditions (**Fig. 6**). These results confirm that Pol II stalling is dynamically regulated and suggest that one of its functions is to prepare genes for activation.

Our genome-wide analysis showed that genes in *Drosophila* embryos are found in three distinct dynamic states: active, stalled or no Pol II. Stalled Pol II is particularly associated with developmental genes that are repressed and poised for activation. We propose that Pol II stalling prepares genes for rapid response to developmental signals during embryogenesis and thus may represent a key regulatory step for gene transcription in development.

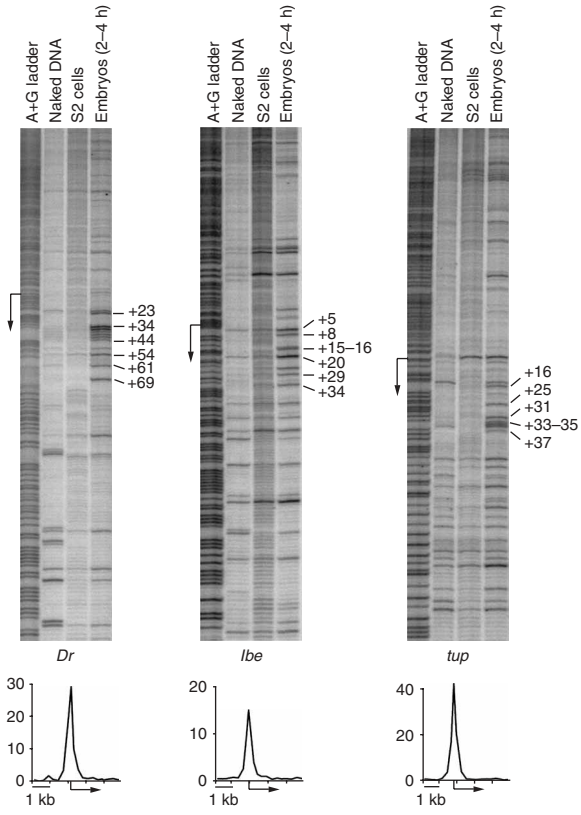


**Figure 6** Pol II stalling at genes before activation. *In vivo* permanganate footprints were carried out on the genes *Dr* and ladybird *(lbe)*, which show a strong 'stalled Pol II' profile in ChIP assays (lower panel) and are activated at later stages. Results for *tup,* a gene known to be repressed at that stage, are shown for comparison. The permanganate footprints were carried out on early wild-type embryos (2–4 h) and S2 cells (107 cells), a cell line derived from older embryos. Positions of A+G (lane 1) and the transcription start site (TSS, arrow) are shown for orientation. T residues sensitive to $KMnO_4$ treatment are shown for naked DNA as control (lane 2), S2 cells (lane 3) and early embryos (lane 4). Hypersensitive T residues in the early embryo sample (selected positions marked on the right) indicate an open transcription bubble. The footprints found at the muscle regulatory genes *Dr* and *lbe* are similar to that seen for *tup*. Thus, Pol II is found in a stalled form at the *Dr* and *lbe* genes before activation during embryonic development.

## METHODS

**Drosophila stock.** *Toll*$^{10b}$ is a dominant gain-of-function mutation of the maternal gene *Toll* (ref. 17). Embryos were collected from *Toll*$^{10b/+}$ females obtained directly from the balanced stock (*Toll*$^{10b}$*/TM3 Sb Ser* and *Toll*$^{10b}$*/OR60*). *Toll*$^{rm9}$ and *Toll*$^{rm10}$ are recessive *Toll* mutations[17], and embryos were collected from *Toll*$^{rm9/rm10}$ females. *gd*$^{7}$ embryos were collected from *gd*$^{7/7}$ females[27]. Wild-type embryos were *white*$^{1118}$.

**ChIP-chip assays.** The chromatin immunoprecipitation (ChIP) experiments coupled to microarrays (chip) were carried out as described[23].

**Antibodies.** The antibodies against Pol II were 8WG16 and H14 (see **Supplementary Note**).

**Arrays.** We used *Drosophila* whole-genome tiling arrays (Agilent) as described[23]. Probes of 60-mers span the entire eukaryotic portion of the *Drosophila melanogaster* genome. Although the spacing of these probes is ~280 bp on average, an additional probe is present between the two probes that flank each known TSS. Thus, the resolution around transcriptional start sites is ~140 bp. The data are available from ArrayExpress and our website, http://web.wi.mit.edu/young/pol2/.

**Data analysis.** We used the Rosetta error model to control for noise at probes; thus a probe required a $P$ value $<0.001$. We did not use our previous algorithms for detecting bound probes and then assigning genes. Rather, we calculated parameters indicating Pol II enrichment directly for each gene (see **Supplementary Note**). A combination of Pol II enrichment at the start site and median enrichment across the gene were used to classify Pol II as either absent, stalled or active.

**Statistics.** For the identification of gene sets that are over-represented in the three classes of genes, we used the hypergeometric distribution test (see **Supplementary Note**).

**Permanganate footprint assays.** We carried out transcription bubble assays with $KMnO_4$ as described previously[4,12]. Embryos were collected 2–4 h after egg deposition (AED), dechorionated and partially homogenized before treatment with $KMnO_4$. Embryos were treated with 20 mM or 40 mM $KMnO_4$ for 60 s on ice. The transcription start sites of the examined genes were identified and confirmed using ESTs in FlyBase and previous expression analysis using tiling arrays[22]. The linker primers and gene-specific primers used for ligation-mediated PCR are listed in the **Supplementary Note**.

**Accession numbers.** ArrayExpress: Data have been deposited with accession code E-TABM-322.

*Note: Supplementary information is available on the Nature Genetics website.*


### ACKNOWLEDGMENTS

We would like to thank R. Zinzen for collecting the *Toll*$^{10b}$, *Toll*$^{rm9/rm10}$ and *gd*$^{7}$ embryos, T. Volkert and J. Love for microarray experimental support and members of the Young laboratory for critical reading of the manuscript. This research was supported in part by the Intramural Research Program of the US National Institutes of Health, National Institute of Environmental Health Sciences (K.A.), US National Institutes of Health grants HG002668 and GM069676 to R.A.Y. and GM34431 to M.L., a grant by the Moore Foundation and a postdoctoral fellowship by the Human Frontier Science Program Organization (HFSPO, A.S.).


### AUTHOR CONTRIBUTIONS

J.Z. and M.L. designed the experiment. J.Z. designed the arrays and carried out the experiments and analysis. A.S., M.K. and J.Z. analyzed expression data and functional categories. J.-W.H., S.N. and K.A. carried out the permanganate footprint assays. J.Z., M.L. and R.A.Y. prepared the manuscript.

### COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics/.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions


1. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
2. Conaway, J.W., Shilatifard, A., Dvir, A. & Conaway, R.C. Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* **25**, 375–380 (2000).
3. Saunders, A., Core, L.J. & Lis, J.T. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**, 557–567 (2006).
4. Gilmour, D.S. & Lis, J.T. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. *Mol. Cell. Biol.* **6**, 3984–3989 (1986).
5. Bender, T.P., Thompson, C.B. & Kuehl, W.M. Differential expression of c-myb mRNA in murine B lymphomas by a block to transcription elongation. *Science* **237**, 1473–1476 (1987).
6. Strobl, L.J. & Eick, D. Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. *EMBO J.* **11**, 3307–3314 (1992).
7. Krumm, A., Meulia, T., Brunvand, M. & Groudine, M. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. *Genes Dev.* **6**, 2201–2213 (1992).
8. Laspia, M.F., Wendel, P. & Mathews, M.B. HIV-1 Tat overcomes inefficient transcriptional elongation in vitro. *J. Mol. Biol.* **232**, 732–746 (1993).
9. Cheng, C. & Sharp, P.A. RNA polymerase II accumulation in the promoter-proximal region of the dihydrofolate reductase and gamma-actin genes. *Mol. Cell. Biol.* **23**, 1961–1967 (2003).
10. Aida, M. *et al.* Transcriptional pausing caused by NELF plays a dual role in regulating immediate-early expression of the junB gene. *Mol. Cell. Biol.* **26**, 6094–6104 (2006).
11. Radonjic, M. *et al.* Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon *S. cerevisiae* stationary phase exit. *Mol. Cell* **18**, 171–183 (2005).
12. Wang, X., Lee, C., Gilmour, D.S. & Gergen, J.P. Transcription elongation controls cell fate specification in the *Drosophila* embryo. *Genes Dev.* **21**, 1031–1036 (2007).
13. Lis, J. & Wu, C. Protein traffic on the heat shock promoter: parking, stalling, and trucking along. *Cell* **74**, 1–4 (1993).
14. Reppas, N.B., Wade, J.T., Church, G.M. & Struhl, K. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Mol. Cell* **24**, 747–757 (2006).
15. Agalioti, T. *et al.* Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* **103**, 667–678 (2000).
16. Soutoglou, E. & Talianidis, I. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* **295**, 1901–1904 (2002).
17. Schneider, D.S., Hudson, K.L., Lin, T.Y. & Anderson, K.V. Dominant and recessive mutations define functional domains of Toll, a transmembrane protein required for dorsal-ventral polarity in the *Drosophila* embryo. *Genes Dev.* **5**, 797–807 (1991).
18. Leptin, M. twist and snail as positive and negative regulators during *Drosophila* mesoderm development. *Genes Dev.* **5**, 1568–1576 (1991).
19. Kosman, D., Ip, Y.T., Levine, M. & Arora, K. Establishment of the mesoderm-neuroectoderm boundary in the *Drosophila* embryo. *Science* **254**, 118–122 (1991).
20. Furlong, E.E., Andersen, E.C., Null, B., White, K.P. & Scott, M.P. Patterns of gene expression during *Drosophila* mesoderm development. *Science* **293**, 1629–1633 (2001).
21. Stathopoulos, A., Van Drenth, M., Erives, A., Markstein, M. & Levine, M. Whole-genome analysis of dorsal-ventral patterning in the *Drosophila* embryo. *Cell* **111**, 687–701 (2002).
22. Biemar, F. *et al.* Comprehensive identification of *Drosophila* dorsal-ventral patterning genes using a whole-genome tiling array. *Proc. Natl. Acad. Sci. USA* **103**, 12763–12768 (2006).
23. Zeitlinger, J. *et al.* Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* **21**, 385–390 (2007).
24. Markstein, M., Markstein, P., Markstein, V. & Levine, M.S. Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. USA* **99**, 763–768 (2002).
25. Tomancak, P. *et al.* Systematic determination of patterns of gene expression during Drosophila embryogenesis. *Genome Biol.* **3**, research0088.1–research0088.14 (2002).
26. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
27. Huang, A.M., Rusch, J. & Levine, M. An anteroposterior Dorsal gradient in the *Drosophila* embryo. *Genes Dev.* **11**, 1963–1973 (1997).
28. Boulay, J.L., Dennefeld, C. & Alberga, A. The *Drosophila* developmental gene snail encodes a protein with nucleic acid binding fingers. *Nature* **330**, 395–398 (1987).
29. Sandmann, T. *et al.* A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. *Dev. Cell* **10**, 797–807 (2006).
30. Furlong, E.E. Integrating transcriptional and signalling networks during muscle development. *Curr. Opin. Genet. Dev.* **14**, 343–350 (2004).