

23. R. Menzel, in *Experimental Behavioral Ecology and Sociobiology*, B. Holldobler, M. Lindauer, Eds. (Sinauer Associates, Sunderland, MA, 1985), pp. 55–74.
24. A. T. Pietrewicz, A. C. Kamil, *Science* **204**, 1332 (1979).
25. B. Z. Kacsóh, T. A. Schlenke, *PLoS ONE* **7**, e34721 (2012).
26. P. H. Harvey, A. Purvis, *Nature* **351**, 619 (1991).

**Acknowledgments:** We thank P. Shen for NPF antiserum; P. Shen, K. Moberg, and S. Sanyal for *Drosophila* strains; and J. van Alphen and B. Wertheim for wasp strains. This work was supported by NIH grant AI081879 to T.A.S. and the Integrated Cellular Imaging Microscopy Core of the Emory Neuroscience NINDS Core Facilities grant, P30NS055077. Data are deposited in the Dryad Repository: <http://dx.doi.org/10.5061/dryad.j5g7m>.

### Supplementary Materials

[www.sciencemag.org/cgi/content/full/339/6122/947/DC1](http://www.sciencemag.org/cgi/content/full/339/6122/947/DC1)  
Materials and Methods  
Figs. S1 to S5  
Tables S1 to S5

3 September 2012; accepted 10 January 2013  
10.1126/science.1229625

# Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing

Hojoong Kwak, Nicholas J. Fuda, Leighton J. Core,\* John T. Lis\*

Transcription regulation occurs frequently through promoter-associated pausing of RNA polymerase II (Pol II). We developed a precision nuclear run-on and sequencing (PRO-seq) assay to map the genome-wide distribution of transcriptionally engaged Pol II at base pair resolution. Pol II accumulates immediately downstream of promoters, at intron-exon junctions that are efficiently used for splicing, and over 3' polyadenylation sites. Focused analyses of promoters reveal that pausing is not fixed relative to initiation sites, nor is it specified directly by the position of a particular core promoter element or the first nucleosome. Core promoter elements function beyond initiation, and when optimally positioned they act collectively to dictate the position and strength of pausing. This "complex interaction" model was tested with insertional mutagenesis of the *Drosophila Hsp70* core promoter.

Tracking the accumulation of RNA polymerase II (Pol II) along genes reveals potential points of regulation (1). For example, a rate-limiting step in early elongation, known as promoter-proximal pausing, has revealed a major regulatory block in the transition to productive elongation in *Drosophila* and mammals (2–8).

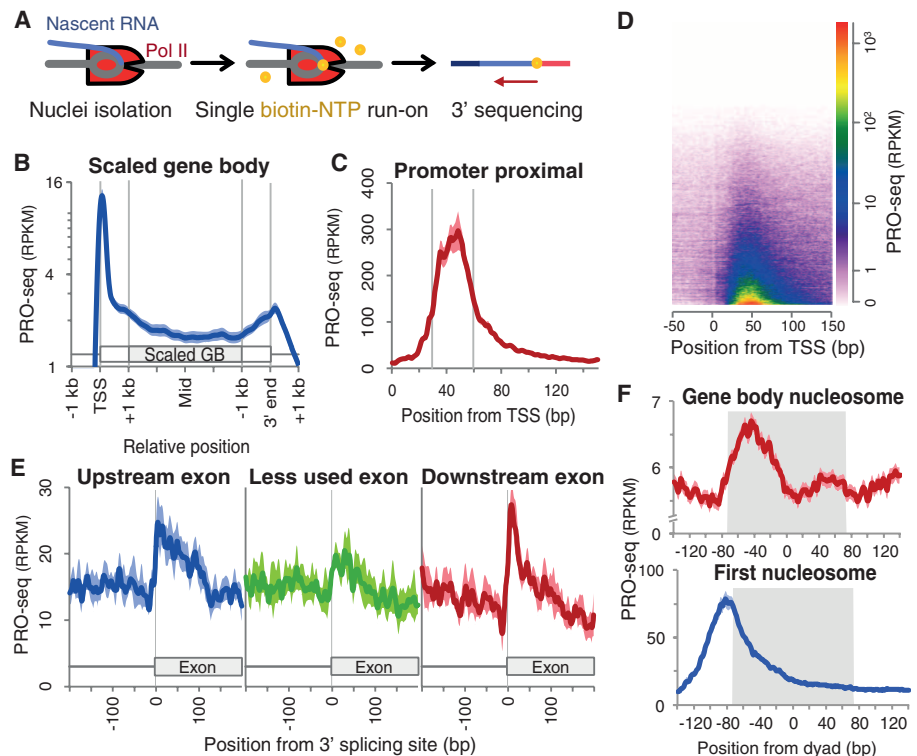
Also, less extensive but substantial accumulation of Pol II over the 3' cleavage/polyadenylation region of genes is proposed to facilitate 3' processing and transcription termination (9, 10). Finally, the interplay of transcription rate and splicing efficiency (11) might be reflected in the selective accumulation of Pol II at splice junctions.

Promoter-associated Pol II pausing is a culmination of intrinsic interactions between Pol II and the underlying DNA, as well as extrinsic stabilization by protein complexes (12). Protein factors such as NELF (negative elongation factor) and DSIF [5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor] (3, 13), DNA elements (14, 15), DNA sequence composition (16), nascent RNA processing (16), and nucleosomes (17) can influence pausing. Understanding how these elements and factors function mechanistically requires a high-resolution view of their spatial relationship. Current tools for precise tracking of the location and status of Pol II in vivo have distinct limitations (18). Chromatin immunoprecipitation-based methods that collect Pol II or associated RNAs do not distinguish paused Pol II from other Pol II–RNA complexes (16, 18, 19). The genome-wide nuclear run-on (GRO-seq) approach (6–8) circumvents these issues by high specific enrichment of nascent transcripts associated with actively engaged

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA.

\*To whom correspondence should be addressed. E-mail: [ljc37@cornell.edu](mailto:ljc37@cornell.edu) (L.J.C.); [jtl10@cornell.edu](mailto:jtl10@cornell.edu) (J.T.L.)

**Fig. 1.** Accumulation of Pol II at promoters, 3' ends, 3' splice sites, and nucleosomes. (A) Schematic of PRO-seq. (B) Average PRO-seq profile of nonoverlapping genes ( $n = 6309$ ) for the sense strand. Gene body (GB) regions (+1 kb from the 5' end to –1 kb from the 3' end) are scaled to 4 kb. Read counts are adjusted to RPKM (reads per kilobase per million mapped reads). Shaded margins surrounding the average plot represent SEM. (C) High-resolution PRO-seq profile from the TSS to +150 bp ( $n = 16,746$ ). (D) Heat map visualization of PRO-seq profile of the annotated genes. Genes are arranged by their increasing PRO-seq density. (E) Average PRO-seq profile at 3' splicing sites of less-used exons and their flanking exons ( $n = 242$  each). Less-used exons have RNA-seq densities less than 5% of their flanking exons (fig. S5B). (F) Average PRO-seq profile relative to the dyad centers of gene body regions and first nucleosomes. The region occupied by nucleosome is shaded gray.



polymerase, but it has a resolution of only 30 to 50 bases (18).

We developed a genome-wide, nuclear run-on assay called PRO-seq that has the sensitivity of GRO-seq but maps Pol II with base pair resolution. PRO-seq uses biotin-labeled ribonucleotide triphosphate analogs (biotin-NTP) for nuclear run-on reactions, allowing the efficient affinity purification of nascent RNAs for high-throughput sequencing from their 3' ends (Fig. 1A and fig. S1A). Supplying only one of the four biotin-NTPs (adenosine, cytosine, guanosine, or uridine triphosphate) restricts Pol II to incorporate a single or at most a few identical bases, resulting in sequence reads that have the same 3' end base within each library (table S1). Moreover, the incorporation of the first biotin base inhibits further transcript elongation, ensuring base pair resolution (fig. S2).

The average profile of PRO-seq density (Fig. 1B) revealed pausing of Pol II immediately

downstream of the transcription start site (TSS) (Fig. 1, C and D) and accumulation of Pol II at 3' cleavage/polyadenylation sites, consistent with previous GRO-seq studies (fig. S3) (20, 21). Interestingly, Pol II also accumulated near 3' splicing sites at spliced exons, but less often at skipped exons (Fig. 1E and fig. S4), which suggests that splicing decisions are connected to differential rates of Pol II elongation through splice junctions (11). Although we have insufficient sequencing coverage to quantify Pol II accumulation at particular 3' splice sites, our composite analyses support a functional coupling between elongation and splicing.

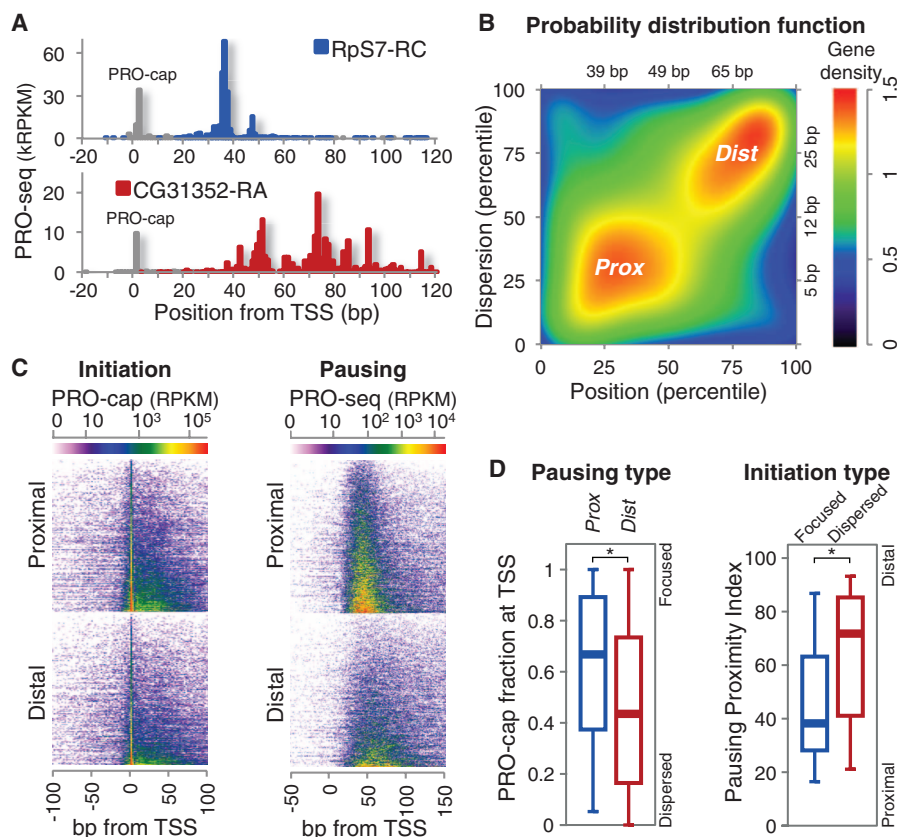
The highest density of PRO-seq reads mapped within positions +30 to +60 from the TSS (Fig. 1, C and D), providing a higher-resolution view of paused Pol II as mapped by GRO-seq (21). Moreover, the pattern of pausing by PRO-seq is consistent with the positions and levels of short nuclear-capped RNAs (scRNAs; fig. S3) (16). Additionally, we

found that PRO-seq maps correspond precisely to positions of engaged Pol II observed in intact cells seen by previous permanganate footprints of transcription bubbles (fig. S2G).

Nucleosomes are known to act as barriers to Pol II (12). In the bodies of genes, the average PRO-seq density showed a relative increase around position  $-40$  from the previously mapped (22) nucleosome centers (Fig. 1F and fig. S5A). This is consistent with measurements of strong DNA-nucleosome interactions (23) and measured impediments to Pol II transcription through nucleosomes measured *in vitro* and in yeast (20, 24). However, the PRO-seq density relative to the first (+1) nucleosome was different (Fig. 1F), with the average PRO-seq density at a maximum around position  $-80$  from the nucleosome centers. Thus, the bulk of promoter-proximal pausing is inconsistent with a standard nucleosome barrier model, at least for *Drosophila*, and is more consistent with tethering of polymerases near the promoter (21).

Although the average promoter-associated pause location is around position +40 from the TSS, pausing is far from uniform. Some genes have more proximal and focused pausing, whereas others have distal and dispersed pausing (Fig. 2A). We systematically assessed genome-wide pausing positions relative to the TSS and their dispersion to identify two characteristic groups of promoters: focused proximal (*Prox*) and dispersed distal (*Dist*) promoters (Fig. 2B and fig. S6). The *Prox* and *Dist* pausing patterns could arise from a fixed length of elongation from initiation sites that have the same dispersion or from variable lengths of elongation from more focused initiation sites. Distinguishing between these possibilities requires precise mapping of the initiation sites, using the same pool of Pol II-engaged nascent RNAs. Therefore, we modified the PRO-seq method to detect initiation sites (PRO-cap; fig. S1B) and compared the degree of variation in the initiation and pause sites. We observed that both *Prox* and *Dist* genes have relatively focused initiation in general (Fig. 2, A and D, and fig. S2G) and that pausing is overall more dispersed than initiation (fig. S7C). Nonetheless, the degree of focused initiation—the fraction of initiation arising at a single TSS—is higher for *Prox* genes, and genes with more focused initiation also have more proximal pausing (Fig. 2D). These findings indicate that although pausing is not fixed to initiation, the mechanisms that produce focused initiation affect the resulting pattern of pausing.

In an effort to otherwise explain the differential patterns of pausing, we first compared the nucleosome occupancy around *Prox* and *Dist* promoters. *Prox* promoters have less nucleosome occupancy than *Dist* promoters (fig. S5B), and some of the Pol II at *Dist* promoters appears to have more intimate contact with the first nucleosome (fig. S5C). These results (and Fig. 1F) support a nucleosome-independent mechanism of pausing for *Prox* promoters, whereas a subset of *Dist* promoters could have a component of pausing

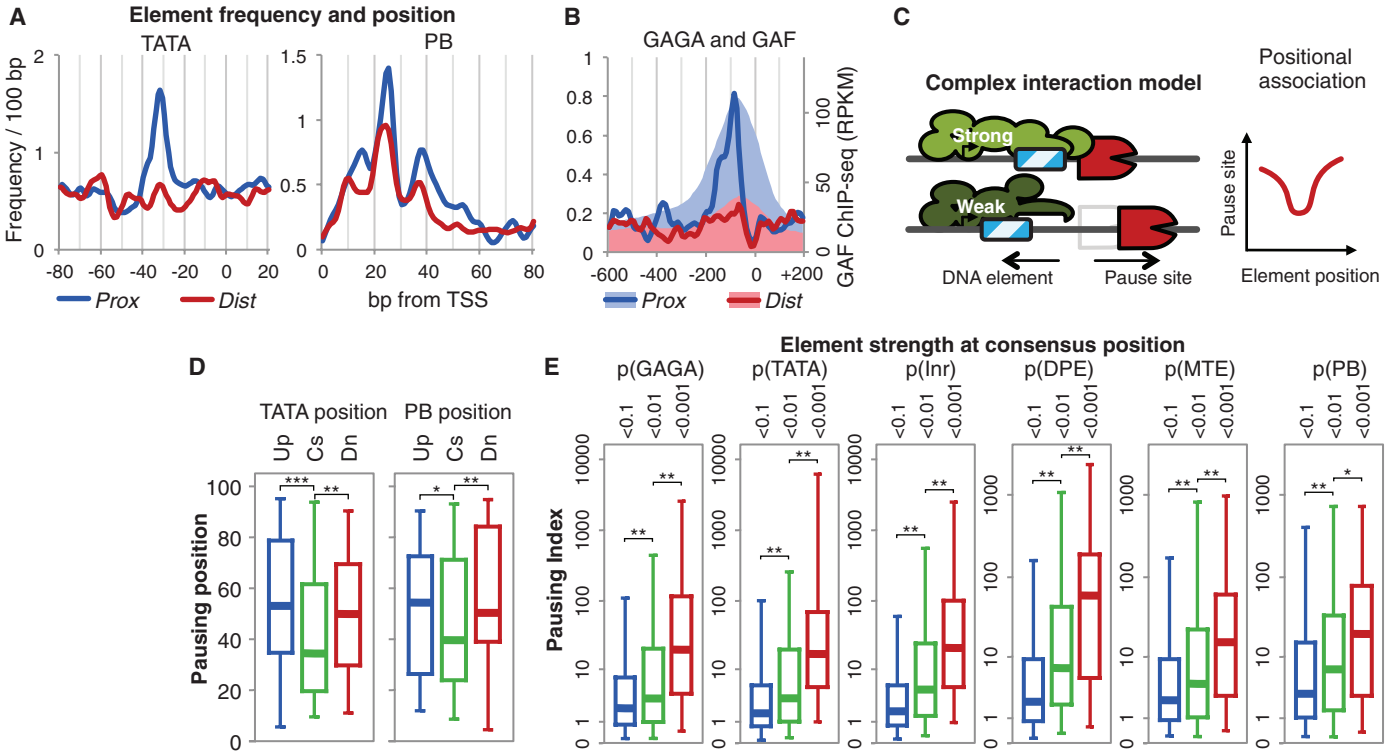


**Fig. 2.** Variations of the pause sites and TSSs. (A) Examples of highly paused genes with different pausing patterns. Initiation sites from PRO-cap mapping are shown in gray. (B) Distribution of paused genes ( $n = 3225$ ) by pausing position and dispersion percentiles. Focused proximal (*Prox*,  $n = 848$ ) and dispersed distal (*Dist*,  $n = 846$ ) groups are indicated; axis units in base pairs are also shown. (C) Heat map of initiation (PRO-cap) and pausing (PRO-seq) for *Prox* and *Dist* genes. (D) Association between initiation and pausing patterns: TSS focusing in *Prox* versus *Dist* genes (left) and pausing proximity in focused versus dispersed initiation genes (right). TSS focusing is measured as the fraction of PRO-cap reads at the TSS ( $\pm 1$  bp) relative to the sum of reads around the TSS ( $\pm 50$  bp) (22). Focused and dispersed initiation genes are the quartiles of the paused genes with the highest and the lowest TSS focusing, respectively. The pausing proximity index is defined by the average of pausing position and dispersion percentiles (fig. S6A). Boxes represent 25th, 50th (median), and 75th percentiles; whiskers are 5th and 95th percentiles. \* $P < 0.001$  [Kolmogorov-Smirnov (KS) test].

that is established by direct nucleosome barriers. Because nucleosome position and occupancy do not explain the bulk of Pol II pausing, we investigated the underlying DNA elements around promoters.

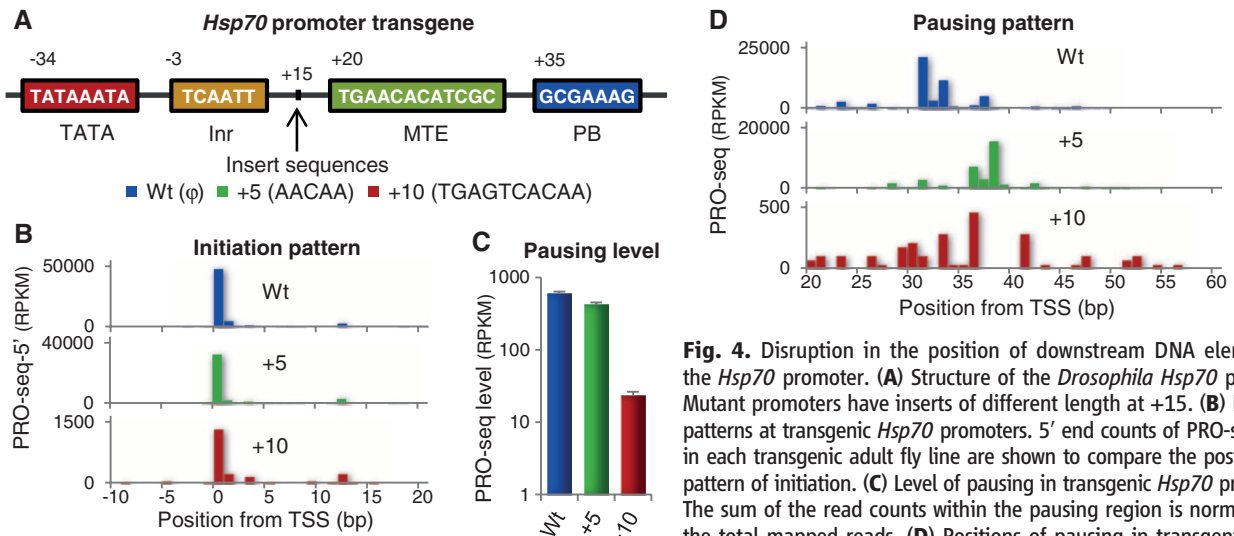
Critical DNA sequence elements within the core promoter direct the position, direction, and efficiency of transcription initiation (25). These include the TATA box, initiator (Inr), motif 10 element (MTE), downstream promoter element

(25), and a recently discovered element implicated in pausing, the pause button (PB) (15) (fig. S8A). Core promoter elements were more enriched on *Prox* promoters than on *Dist* promoters (Fig. 3A and fig. S8, B to D). Additionally, when



**Fig. 3.** Relationship between promoter DNA elements and Pol II pausing. (A) Frequency of TATA box and PB in *Prox* and *Dist* subsets. The average frequency per gene is shown. (B) Frequency of GAGA element (lines) and GAGA-factor binding (shaded areas) (27) in *Prox* and *Dist* subsets. (C) The complex interaction model of relationship between DNA elements and paused Pol II. The DNA elements (blue) are at their consensus (strong) or slightly upstream (weak) positions; the expected changes of the pausing positions are plotted. (D) Pattern of the positional association between

DNA elements and pausing positions. Pausing position percentiles are shown for gene subsets according to element position (Up, upstream; Cs, optimal consensus position; Dn, downstream; subset information in table S2). \* $P < 0.14$ , \*\* $P < 0.06$ , \*\*\* $P < 0.002$  (KS test). (E) Association of promoter DNA element strength at consensus positions with pausing index (6). Active genes ( $n = 5471$ ) are divided into three subsets according to the distance-weighted  $P$  values of the DNA elements to the consensus positions (table S3). \* $P < 0.01$ , \*\* $P < 0.001$  (KS test).



**Fig. 4.** Disruption in the position of downstream DNA elements in the *Hsp70* promoter. (A) Structure of the *Drosophila Hsp70* promoter. Mutant promoters have inserts of different length at +15. (B) Initiation patterns at transgenic *Hsp70* promoters. 5' end counts of PRO-seq reads in each transgenic adult fly line are shown to compare the position and pattern of initiation. (C) Level of pausing in transgenic *Hsp70* promoters. The sum of the read counts within the pausing region is normalized to the total mapped reads. (D) Positions of pausing in transgenic *Hsp70* promoters.



we searched within the extended promoter regions of *Prox* and *Dist* groups for the presence of 232 additional functional DNA elements (26) (fig. S8E), only the GAGA element, residing ~80 base pairs (bp) upstream of the TSS (3, 15), showed strong associations with *Prox* genes (Fig. 3B), as did the level of GAGA-factor binding (27). Thus, core promoter elements and GAGA factor appear to play an important role in the mechanism of pausing.

Pausing positions could be determined through direct tethering of elongating Pol II to promoter elements. Alternatively, in a “complex interaction” model, pausing could be mediated through protein complexes that function best when cognate elements are located at specific positions in the core promoter. Thus, if we examine the association of the positions of the DNA elements and the pausing sites in this model, we expect a V-shaped plot of association rather than a simple linear correlation. Displacement of the element from the optimal position will weaken the interactions within the core complex, resulting in downstream scattering and a reduced level of pausing (Fig. 3C). To test this, we examined genes in which a particular promoter DNA element occurs only once, and divided genes into three subsets: the optimal consensus position, upstream, and downstream. Genes with the DNA elements nearest to the consensus positions showed more proximal pausing. Genes with TATA near -30 had more proximal pausing than genes with TATA at positions of -40 or more, showing a V-shaped association (Fig. 3D). This V pattern was observed in both the upstream elements TATA and Inr (fig. S9A) and the downstream elements PB (Fig. 3D) and MTE (fig. S9B). Also, pausing tended to be stronger in genes with the elements at the optimal positions (fig. S9D). Furthermore, the extent of pausing showed strong dependency on the match of the DNA elements to their consensus sequence and consensus positions (Fig. 3E). Together, these patterns of association between core promoter elements and pausing support the complex interaction model and explain the strong and focused pausing on *Prox* promoters.

The complex interaction model depends on both the presence and the correct positioning of core promoter elements. We disrupted the positional relationship of core elements in the well-studied *Drosophila* gene *Hsp70* (1). Transgenic fly lines were generated that carry mutant *Hsp70* promoters with spacers inserted at the +15 position between the upstream and downstream promoter elements (Fig. 4A) and were analyzed by PRO-seq (fig. S10). The initiation sites remained constant in these mutant promoters, as indicated by the 5' ends of the PRO-seq reads (Fig. 4B). The transgenic *Hsp70* without spacers showed a strong pause peak mainly at position +31 (Fig. 4D). When a 5-bp spacer was inserted, the pause peak was shifted 5 to 7 bp downstream from the original site. Because additional bases were transcribed before pausing, the position of pausing is not predetermined by elongation distance. When

a 10-bp spacer was inserted, pausing sites became scattered between positions +20 and +60 (Fig. 4D) and had fewer reads (Fig. 4C). Collectively, these results support the core interaction model and suggest that the interaction complex can accommodate a small change (5 bp) in the positional context of the DNA sequences, whereas a larger change (10 bp) results in reduced and dispersed pausing.

The advances in resolution provided by PRO-seq enable the precise and genome-wide assessment of the relationship between promoter-proximal pausing and the core promoter structure. For the strong and tightly clustered pausing of the *Prox* genes, our results provide support for a complex interaction model involving the promoter initiation complex, which can extend up to 30 bp from the TSS (28), physically contacting and tethering the pausing complexes. This may share a kinship with bacterial initiation factor  $\sigma$ , which is retained within the early elongation complex and interacts with promoter proximal DNA during transcription pausing in *Escherichia coli* (29). It is noteworthy that the *Prox* genes are expressed on average at a lower level but show a broader range of expression (fig. S6D), and that the *Dist* genes are enriched in constitutively active genes (table S5). These results suggest that the mechanistic distinctions have regulatory consequences. A well-structured core promoter may strongly recruit Pol II; however, it can also effectively retain Pol II in a paused configuration close to the TSS, until activation signals allow its escape into productive elongation.

#### References and Notes

1. N. J. Fuda, M. B. Ardehali, J. T. Lis, *Nature* **461**, 186 (2009).
2. E. B. Rasmussen, J. T. Lis, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7923 (1993).
3. C. Lee *et al.*, *Mol. Cell. Biol.* **28**, 3290 (2008).
4. G. W. Muse *et al.*, *Nat. Genet.* **39**, 1507 (2007).
5. J. Zeitlinger *et al.*, *Nat. Genet.* **39**, 1512 (2007).
6. L. J. Core, J. J. Waterfall, J. T. Lis, *Science* **322**, 1845 (2008).
7. N. Hah *et al.*, *Cell* **145**, 622 (2011).

8. I. M. Min *et al.*, *Genes Dev.* **25**, 742 (2011).
9. L. J. Core, J. T. Lis, *Science* **319**, 1791 (2008).
10. K. Glover-Cutter, S. Kim, J. Espinosa, D. L. Bentley, *Nat. Struct. Mol. Biol.* **15**, 71 (2008).
11. A. R. Kornblitt, *Nat. Struct. Mol. Biol.* **13**, 5 (2006).
12. A. Saunders, L. J. Core, J. T. Lis, *Nat. Rev. Mol. Cell Biol.* **7**, 557 (2006).
13. C. H. Wu *et al.*, *Genes Dev.* **17**, 1402 (2003).
14. C. H. Wu *et al.*, *Mol. Cell. Biol.* **21**, 1593 (2001).
15. D. A. Hendrix, J. W. Hong, J. Zeitlinger, D. S. Rokhsar, M. S. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7762 (2008).
16. S. Nechaev *et al.*, *Science* **327**, 335 (2010).
17. T. N. Mavrich *et al.*, *Nature* **453**, 358 (2008).
18. S. Nechaev, K. Adelman, *Biochim. Biophys. Acta* **1809**, 34 (2011).
19. L. S. Churchman, J. S. Weissman, *Nature* **469**, 368 (2011).
20. E. Larsson *et al.*, *Nature* **471**, 115 (2011).
21. L. J. Core *et al.*, *Cell Rep.* **2**, 1025 (2012).
22. D. A. Gilchrist *et al.*, *Cell* **143**, 540 (2010).
23. M. A. Hall *et al.*, *Nat. Struct. Mol. Biol.* **16**, 124 (2009).
24. V. A. Bondarenko *et al.*, *Mol. Cell* **24**, 469 (2006).
25. T. Juven-Gershon, J. Y. Hsu, J. W. Theisen, J. T. Kadonaga, *Curr. Opin. Cell Biol.* **20**, 253 (2008).
26. A. Stark *et al.*; Harvard FlyBase curators; Berkeley *Drosophila* Genome Project, *Nature* **450**, 219 (2007).
27. S. Sharma, M. J. Guertin, J. T. Lis, undergraduate thesis, Cornell University, Ithaca, NY (2012).
28. P. A. Emanuel, D. S. Gilmour, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8449 (1993).
29. B. Z. Ring, W. S. Yarnell, J. W. Roberts, *Cell* **86**, 485 (1996).

**Acknowledgments:** Supported by NIH grants GM25232 and HG004845 (J.T.L.) and a Howard Hughes Medical Institute fellowship (H.K.). Sequence data are in the Gene Expression Omnibus (GEO) database under accession number GSE42117. Part of this work is included in a broader U.S. patent application (12/554,472, “Genome-wide Method for Mapping of Engaged RNA Polymerases Quantitatively and at High Resolution”), which refers to variants of the method that we use here.

#### Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6122/950/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S10  
Tables S1 to S5  
Source Codes for Analysis Scripts  
References

27 August 2012; accepted 21 December 2012  
10.1126/science.1229386

## Little Emperors: Behavioral Impacts of China's One-Child Policy

L. Cameron,<sup>1\*</sup> N. Erkal,<sup>2</sup> L. Gangadharan,<sup>3</sup> X. Meng<sup>4</sup>

We document that China's One-Child Policy (OCP), one of the most radical approaches to limiting population growth, has produced significantly less trusting, less trustworthy, more risk-averse, less competitive, more pessimistic, and less conscientious individuals. Our data were collected from economics experiments conducted with 421 individuals born just before and just after the OCP's introduction in 1979. Surveys to elicit personality traits were also used. We used the exogenous imposition of the OCP to identify the causal impact of being an only child, net of family background effects. The OCP thus has significant ramifications for Chinese society.

China's One-Child Policy (OCP) restricts the number of children that urban couples can have to one, with exceptions for those from ethnic minorities or with a severely disabled child. The policy has given rise to a land of “little emperors,” whose parents dote on them exclu-

sively. This has led to widespread concern within China about the social skills of this generation and the observation that these children tend to be more self-centered and less cooperative (1–3). This can be seen in developments such as employers including phrases like “no single children”



**Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing**

Hojoong Kwak, Nicholas J. Fuda, Leighton J. Core and John T. Lis  
(February 21, 2013)  
*Science* **339** (6122), 950-953. [doi: 10.1126/science.1229386]

Editor's Summary

**Pressing Pause**

The pausing of RNA polymerase II (Pol II) near promoters is an important and distinct regulatory step in the process of transcribing DNA into RNA. Although protein factors and DNA elements involved in pausing have been identified, the post-initiation events that lead to pausing are poorly understood. **Kwak *et al.*** (p. 950) present an approach, termed PRO-seq, for monitoring the position of actively transcribing RNA polymerase with single-nucleotide precision. The approach was used to explore the mechanism of promoter proximal pausing and showed that precise positioning between core promoter elements and the pause site is critical. The findings show how the promoter dictates transcriptional pausing and detects the preferential localization of active transcription complexes within the genome.

---

This copy is for your personal, non-commercial use only.

---

**Article Tools** Visit the online version of this article to access the personalization and article tools:  
<http://science.sciencemag.org/content/339/6122/950>

**Permissions** Obtain information about reproducing this article:  
<http://www.sciencemag.org/about/permissions.dtl>

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.