Please cite this article in press as: Maxwell et al., Pol II Docking and Pausing at Growth and Stress Genes in *C. elegans*, Cell Reports (2014), http://dx.doi.org/10.1016/j.celrep.2014.01.008

Cell Reports
Article

Pol II Docking and Pausing at Growth and Stress Genes in *C. elegans*

Colin S. Maxwell,[1] William S. Kruesi,[2] Leighton J. Core,[3] Nicole Kurhanewicz,[1,5] Colin T. Waters,[3] Caitlin L. Lewarch,[1,6] Igor Antoshechkin,[4] John T. Lis,[3] Barbara J. Meyer,[2] and L. Ryan Baugh[1,*]
[1]Department of Biology, Duke Center for Systems Biology, Duke University, Durham, NC 27708, USA
[2]Howard Hughes Medical Institute, Department of Molecular and Cellular Biology, University of California, Berkeley, Berkeley, CA 94720, USA
[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA
[4]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA
[5]Present address: 130 Mason Farm Road, 1125 Bioinformatics Building, CB# 7108, University of North Carolina, Chapel Hill, NC 27599, USA
[6]Present address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
*Correspondence: ryan.baugh@duke.edu
http://dx.doi.org/10.1016/j.celrep.2014.01.008

SUMMARY

Fluctuations in nutrient availability profoundly impact gene expression. Previous work revealed postrecruitment regulation of RNA polymerase II (Pol II) during starvation and recovery in *Caenorhabditis elegans*, suggesting that promoter-proximal pausing promotes rapid response to feeding. To test this hypothesis, we measured Pol II elongation genome wide by two complementary approaches and analyzed elongation in conjunction with Pol II binding and expression. We confirmed bona fide pausing during starvation and also discovered Pol II docking. Pausing occurs at active stress-response genes that become downregulated in response to feeding. In contrast, "docked" Pol II accumulates without initiating upstream of inactive growth genes that become rapidly upregulated upon feeding. Beyond differences in function and expression, these two sets of genes have different core promoter motifs, suggesting alternative transcriptional machinery. Our work suggests that growth and stress genes are both regulated postrecruitment during starvation but at initiation and elongation, respectively, coordinating gene expression with nutrient availability.

INTRODUCTION

All organisms must cope with fluctuations in environmental conditions. There is a pervasive difference in the genes expressed during stressful conditions and those that support growth, demonstrating a fundamental role of transcriptional regulation (Gasch et al., 2000). Rapid and coordinated responses to changes in environmental conditions are essential, but the mechanisms responsible are not well understood.

For the nematode *C. elegans*, life in the wild is feast or famine, making it an ideal metazoan model to investigate transcriptional responses to nutrient availability. Larvae that hatch without food arrest development in the first larval stage (L1 arrest or L1 diapause) and become resistant to stress (Baugh, 2013). Arrested larvae respond rapidly to feeding, dramatically altering gene expression and initiating growth (Baugh et al., 2009; Maxwell et al., 2012). During L1 arrest, RNA polymerase II (Pol II) accumulates at the 5′ end of genes that are upregulated during recovery (Baugh et al., 2009), suggesting that postrecruitment regulation of Pol II contributes to nutritional control of transcription.

It has become clear in recent years that postrecruitment regulation of early elongation (pausing) is widespread in the animals where it has been investigated (Core et al., 2008; Kim et al., 2005; Muse et al., 2007; Rahl et al., 2010; Zeitlinger et al., 2007). Pausing has been suggested to promote rapid response to changes in environmental conditions and during development, as in the heat shock response where it was first discovered (Muse et al., 2007; Rougvie and Lis, 1988; Zeitlinger et al., 2007). However, pausing does not always predict upregulation in models of inducible gene expression (Gilchrist et al., 2012; Hah et al., 2011; Lin et al., 2011).

We previously used Pol II chromatin immunoprecipitation and sequencing (ChIP-seq) to show that the polymerase accumulates at the 5′ end of many genes during L1 arrest (Baugh et al., 2009). We hypothesized that this accumulation reflects Pol II pausing. However, a "paused" polymerase is defined as having initiated elongation but transiently halted (Adelman and Lis, 2012), and ChIP cannot distinguish between elongating and nonelongating Pol II. In addition, the multimeric negative elongation factor (NELF) contributes to pausing in other systems (Nechaev and Adelman, 2011; Renner et al., 2001; Wu et al., 2003), but none of its subunits has homologs in the *C. elegans* genome (Narita et al., 2003). Furthermore, *trans*-splicing obscures the transcription start site (TSS) of most genes in *C. elegans* (Allen et al., 2011), making interpretation of Pol II accumulation difficult.

Our results here suggest that two independent forms of postrecruitment regulation occur during starvation in *C. elegans*, docking and pausing, affecting growth and stress genes,

Please cite this article in press as: Maxwell et al., Pol II Docking and Pausing at Growth and Stress Genes in *C. elegans*, Cell Reports (2014), http://dx.doi.org/10.1016/j.celrep.2014.01.008

Cell PRESS
Open ACCESS

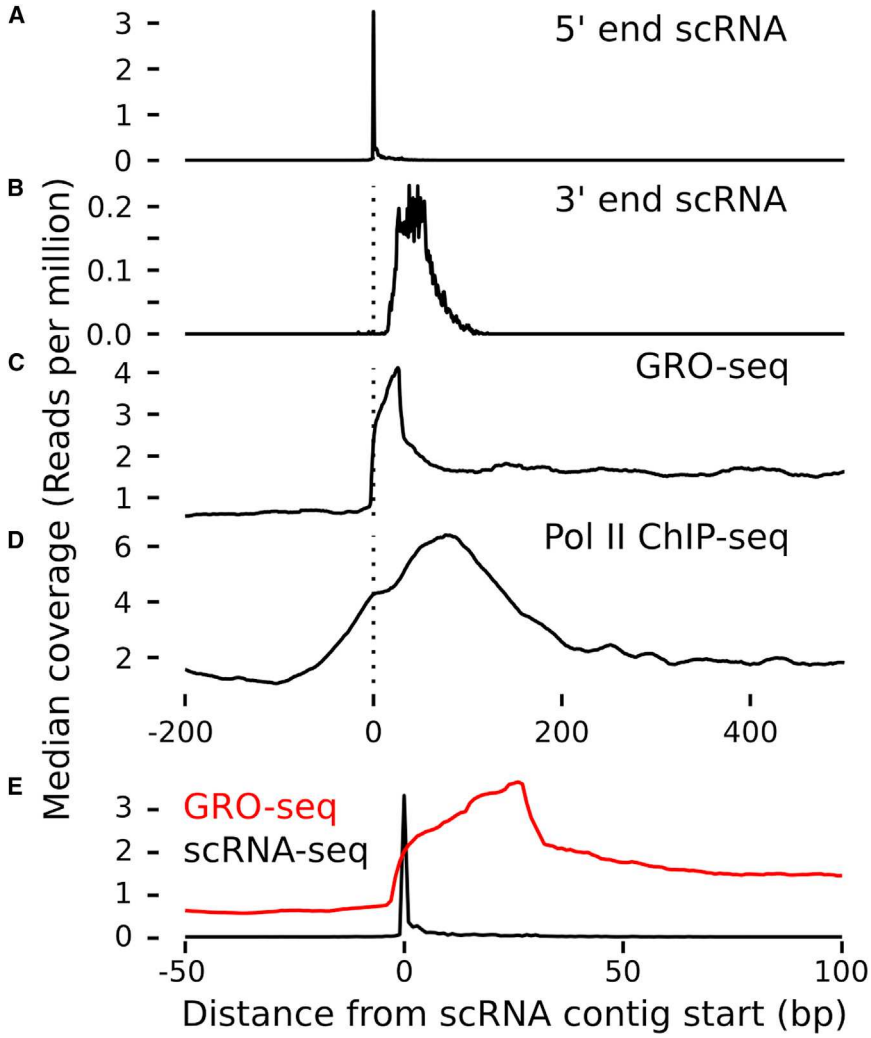

Figure 1. scRNAs Coincide with Accumulation of Elongating RNA Pol II

(A–D) Coverage of (A) the 5′ end of scRNAs, (B) the 3′ end of scRNAs, (C) GRO-seq reads, and (D) Pol II ChIP-seq reads is plotted relative to the beginning of 789 contig regions of scRNA coverage. (E) Coverage of 5′ end of scRNA reads (black) and GRO-seq reads (red) in the immediate proximity of the contig start is plotted. Only scRNA contigs within 100 bp of annotated TSSs for protein-coding genes are included (WS220). Coverages are the median bootstrap estimates of the mean. See also Figures S1–S3.

respectively. Integrated analysis of Pol II binding, nascent transcript production, and elongation confirms that Pol II pausing occurs in starved *C. elegans* L1 stage larvae, where it is associated with active stress-response genes. Furthermore, a TFIIS mutant suggests backtracking of paused polymerase as in other systems. Surprisingly, this analysis also revealed that "docked" Pol II accumulates without initiating transcription just upstream of TSSs of growth genes. In addition to encoding proteins with distinct functions, genes associated with docking and pausing respond in opposite ways to feeding and are enriched for different core promoter motifs. Our results reveal a fundamental distinction between growth and stress genes and suggest that this difference extends to mechanisms of postrecruitment transcriptional regulation.

RESULTS

Nascent RNA Sequencing by scRNA-Seq and GRO-Seq

Our published Pol II ChIP-seq analysis could not distinguish between inactive and elongating polymerase (Baugh et al., 2009). To address this, we sequenced short, capped RNAs (scRNA-seq) to measure elongation activity genome wide. Nascent RNAs are capped on their 5′ end in *Drosophila* (Rasmussen and Lis, 1993), providing a strategy to clone and sequence them as scRNA (Nechaev et al., 2010). In addition, sequencing the 3′ end of scRNA reveals the location of promoter-proximal Pol II with nucleotide (nt) resolution (Nechaev et al., 2010). We performed a variety of control experiments that demonstrate the sensitivity, specificity, reproducibility, and fidelity of our scRNA-seq procedure (Figures S1 and S2; Supplemental Information). These control experiments show that we are able to specifically detect nascent elongation products as scRNAs in *C. elegans*.

We also analyzed nascent RNAs using global nuclear run-on sequencing (GRO-seq) data as an independent measurement of elongating Pol II (Kruesi et al., 2013). Unlike scRNA-seq, which only reports on elongation activity near the TSS, GRO-seq reports on elongation throughout the gene. scRNA-seq also cannot distinguish between RNA species that remain associated with paused Pol II and those that have been released through termination of transcription, but GRO-seq can (Adelman and Lis, 2012). However, GRO-seq does not locate the position of paused polymerase with the single-nt resolution of 3′ scRNA-seq. For these reasons, scRNA-seq and GRO-seq provide complementary ways to interrogate elongation.

scRNA-Seq Reveals Pol II Pausing in *C. elegans*

scRNA detection coincides with Pol II accumulation consistent with promoter-proximal pausing. We used scRNA-seq on a pair of biological replicates starved during the L1 larval stage. We identified 789 contiguous regions of 5′ scRNA-seq coverage (scRNA contigs) that are within 100 bp of the annotated TSS of protein-coding genes. We found that these scRNA contigs are associated with accumulation of Pol II detected by ChIP-seq and GRO-seq (Figure 1), suggesting pausing of the polymerase

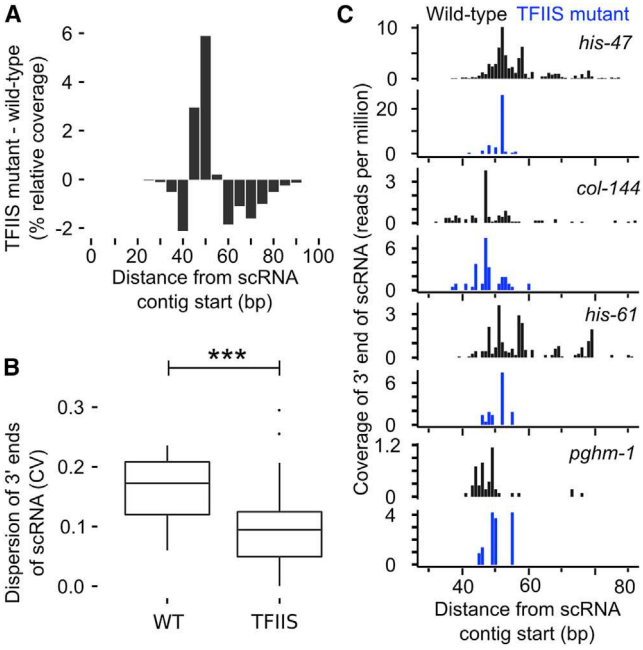

Figure 2. TFIIS Mutation Alters the Size Distribution of scRNAs

(A) The difference in relative coverage between wild-type and the TFIIS mutant is plotted relative to the beginning of 789 scRNA contigs within 100 bp of annotated protein-coding gene TSSs. Each bin shows the mean change in coverage over a 5 bp window. Coverage is the median bootstrap estimate of the mean.

(B) A boxplot comparing the CV for the distance between the 3′ ends of scRNAs and the beginning of the contig is plotted. The CV was calculated for each contig separately. In order to address possible effects resulting from the smaller TFIIS mutant library sizes, the wild-type data were resampled to calculate a bootstrap estimate of the CV.

(C) Four examples of the distribution of the 3′ end of scRNAs are plotted for wild-type and the TFIIS mutant. All genes are plotted with their 5′ end to the left regardless of strand.

during early elongation. Unless stated otherwise, the Pol II ChIP-seq data presented are from Zhong et al. (2010) (Table S1). The 5′ end of scRNA contigs and GRO-seq signal align precisely (Figure 1E), demonstrating the precision and consistency of these two data sets. These results also suggest that scRNA synthesis initiates at the same position as mRNA synthesis, consistent with scRNAs being nascent transcription products.

Pol II pauses in approximately the same location relative to the TSS in *C. elegans* as other animals. Sequencing the 5′ and 3′ ends of scRNAs allows us to determine their size distribution, revealing the distance traveled by Pol II prior to pausing. Consistent with *Drosophila* (Nechaev et al., 2010), 75% of scRNAs are 25–65 nt long (Figure 1B). There is also good agreement between the position of the 3′ ends of scRNA and the peak of Pol II accumulation based on ChIP-seq, as expected for pausing (Figure 1). The amount of scRNA correlates with Pol II ChIP-seq and GRO-seq coverage over scRNA contig coordinates (Spearman's r = 0.37 and 0.30, respectively; Spearman's rank correlation test, $p < 2 \times 10^{-16}$ for both comparisons), suggesting that these assays detect the same population of paused Pol II molecules. Together, these data provide evidence that Pol II is paused in the promoter-proximal region of many genes during L1 arrest (see below for estimation of the number of paused genes).

TFIIS and Backtracking of Paused Pol II

We found that paused Pol II is prone to backtracking in *C. elegans*. When Pol II pauses in other organisms, it can backtrack a few base pairs, and the general transcription factor TFIIS helps it resume elongation (Adelman et al., 2005; Kettenberger et al., 2003). Depleting TFIIS in yeast or *Drosophila* S2 cells results in net elongation of nascent RNAs near pause sites, reflecting backtracking without cleavage (Churchman and Weissman, 2011; Nechaev et al., 2010). To examine the function of TFIIS in *C. elegans*, we sequenced the 3′ end of scRNAs in a TFIIS mutant (T24H10.1(*ok2749*)) during L1 arrest. The mutant has a significantly different scRNA size distribution (Kolmogorov-Smirnov test, $p < 2.2 \times 10^{-16}$). In particular, fewer of the shortest scRNAs and more of the moderately sized scRNAs are detected in the TFIIS mutant (Figure 2A). This increase in scRNA length is consistent with TFIIS relieving backtracking. Furthermore, these results provide strong evidence that the scRNAs we detect are nascent transcription products as opposed to degradation products.

TFIIS also affects the dispersion of pause sites in individual genes. We found that the longest scRNAs are actually less abundant in the TFIIS mutant than wild-type (Figure 2A). This result suggests that TFIIS is required for Pol II to escape relatively proximal pause sites and reach secondary pause sites where it is associated with longer scRNAs. To address this hypothesis, we examined the dispersion of pause sites in individual genes. We calculated the coefficient of variation (CV) for the distance between the 3′ ends of scRNAs and the start of the contig within individual contigs. This analysis revealed that there is a smaller CV in the TFIIS mutant than in wild-type (Figure 2B; Wilcoxon signed rank test, $p = 5.2 \times 10^{-13}$). This observation is consistent with the mutant having a relatively narrow scRNA size distribution genome wide (Figure 2A), but it shows that the effect occurs at individual loci. Examination of 3′ scRNA ends at individual genes supports this interpretation (Figure 2C). These results suggest that Pol II pauses in a more focused region in the TFIIS mutant than wild-type, as if TFIIS helps Pol II escape proximal pause sites in order to pause at more distal sites.

TSS Identification and the Frequency of Pausing

The majority of mRNA transcripts in *C. elegans* have a 22 nt leader sequence added to their 5′ end in a cotranscriptional *trans*-splicing reaction (Allen et al., 2011). As a result, current genome annotation of TSSs actually corresponds to *trans*-splice sites in the majority of cases. Given that GRO-seq detects nascent RNAs, it provides an opportunity to discover true TSSs. To increase coverage of GRO-seq signal at 5′ ends, we used data from a variant of GRO-seq called GRO-cap to sequence only the capped 5′ end of nascent transcripts. Development of this technique, validation, results, and criteria used to identify TSSs are presented elsewhere (Kruesi et al., 2013). We pooled TSS calls generated from embryos, L1 arrest, and fed L3 stage larvae to generate 5,192 high-confidence true TSSs for protein-coding genes, excluding TSSs found inside operons.

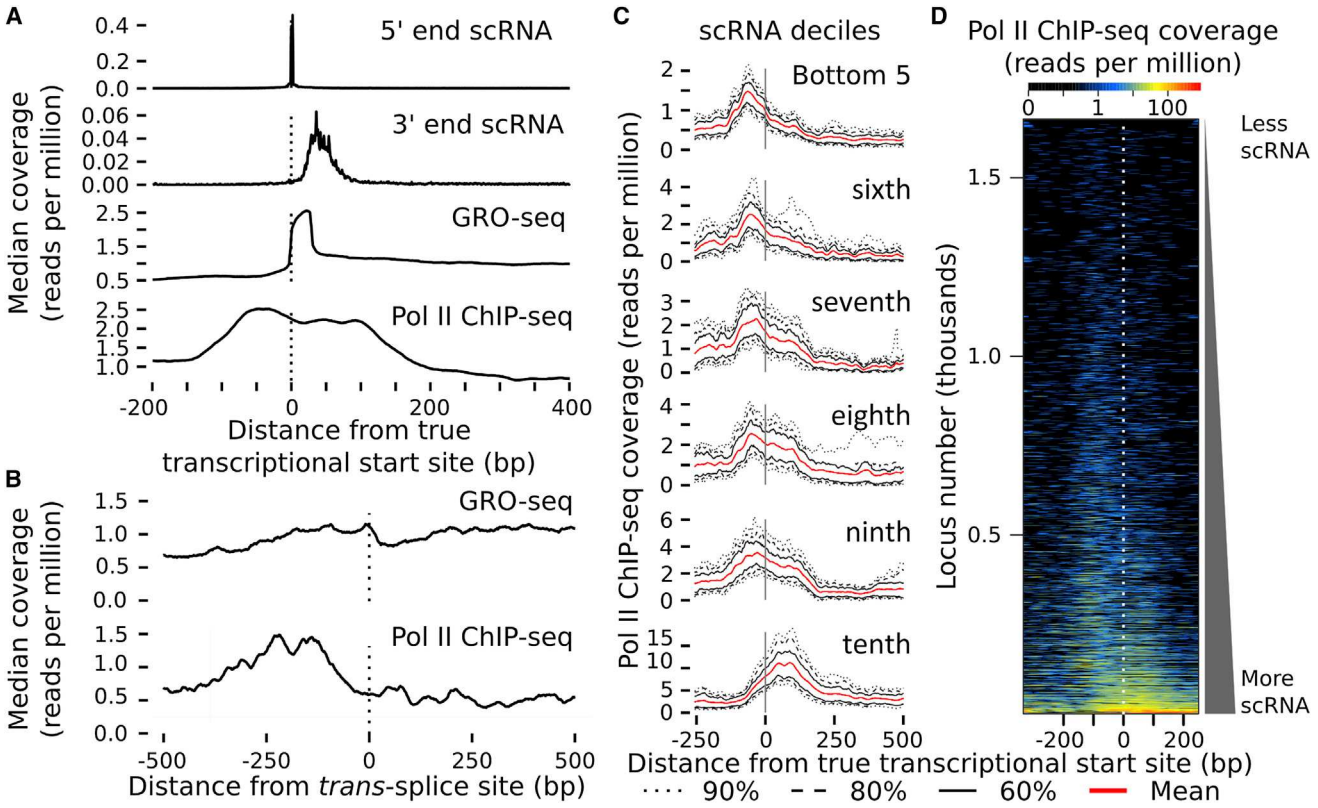

Figure 3. Pol II Accumulates Upstream of True TSSs at Genes with Relatively Little Elongation

(A) Coverage of the 5′ end of scRNAs, 3′ end of scRNAs, GRO-seq reads, and Pol II ChIP-seq reads is plotted relative to 5,192 true TSSs.

(B) Coverage of GRO-seq and Pol II ChIP-seq reads is plotted relative to 590 empirically identified SL1 *trans*-splice sites. Coverages for (A) and (B) are the median bootstrap estimates of the mean.

(C) Mean Pol II ChIP-seq coverage around 5,192 true TSSs is plotted for deciles of scRNA abundance mapping within 100 bp downstream of the TSS. The bottom five deciles of scRNA are each made up of loci with no scRNAs mapping to them and are merged. Dotted, dashed, and solid black lines show the 90%, 80%, and 60% bootstrap confidence intervals of the mean, respectively, based on computing the mean of a sample of 10% of the data 1,000 times. F44E5.4 and F44E5.5 were omitted (see Supplemental Experimental Procedures).

(D) A heatmap of Pol II ChIP-seq coverage is plotted for the same genes as in (C). Genes are sorted by the number of scRNA reads mapping within 100 bp downstream of the TSS.

See also Figures S4 and S5.

We performed scRNA-seq only in L1 arrest, and we detected scRNAs (with a false discovery rate [FDR] of 1%) at 29% of true TSSs. The 5′ end of scRNAs agrees extremely well with true TSSs (Figure 3A). The precise registration of these two data sets reflects the reliability of the TSS calls.

We verified that scRNAs are generated primarily from paused Pol II. Pol II does not appear to pause for the *trans*-splicing reaction based on Pol II ChIP-seq and GRO-seq (Figure 3B). We therefore used the frequency of scRNA reads that begin with the 22 nt SL1 splice leader and extend into the *trans*-spliced gene to assess the relative contribution of elongating Pol II to scRNA abundance. We assume that scRNAs present at the *trans*-splice site are produced exclusively by readthrough of elongating Pol II (because there is no pausing) and that scRNAs at TSSs are produced by a combination of readthrough and pausing. The average ratio of scRNAs at TSSs versus *trans*-splice sites is 4.3, suggesting that on average, about 75% of the scRNAs present at TSSs are due to pausing (Supplemental Information).

scRNA-seq reveals paused Pol II immediately downstream of true TSSs. We applied a cutoff of 20 reads of scRNA at the TSS and found 8.2% of genes with evidence of Pol II pausing by this criterion (Table S2). This frequency is consistent with an independent analysis of the GRO-seq pausing index (Kruesi et al., 2013). However, caveats apply to the interpretation of these results. For example, whole animals were used for all of our measurements, presumably affecting sensitivity, and an arbitrary cutoff is used in both cases. Furthermore, paused Pol II is not the only source of scRNA (see above and the Supplemental Information), and the GRO-seq pause index is very sensitive to weak signal in the body of the gene (the denominator of the index). Caveats aside, it is noteworthy that pausing in starved *C. elegans* larvae appears less common than in other systems where it has been investigated. Nevertheless, this analysis suggests that Pol II pauses downstream of true TSSs during starvation in *C. elegans* (for examples, see Figure S3).

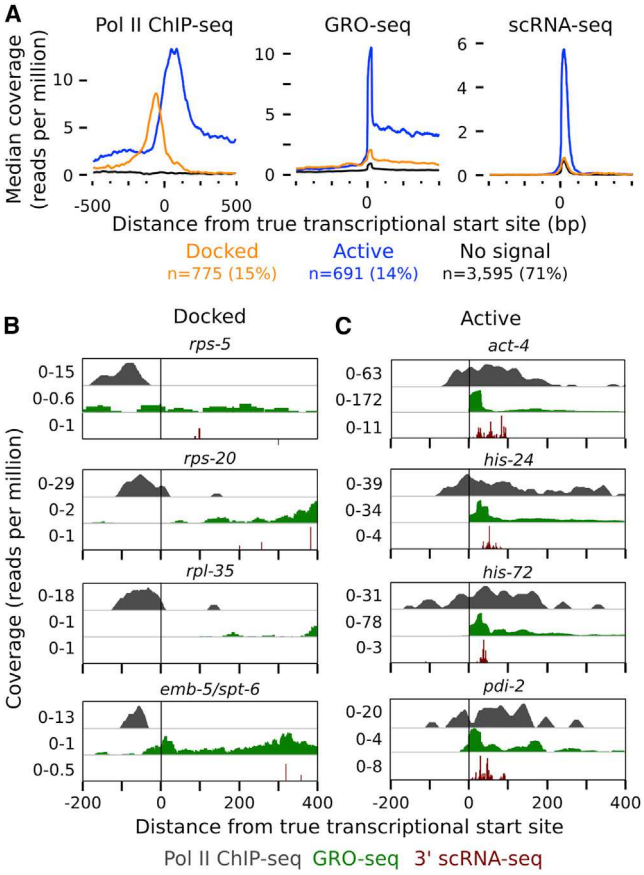

Figure 4. Clustering Genes Based on Patterns of Pol II ChIP-Seq Coverage around True TSSs Identifies Genes with "Docked" and "Active" Pol II

(A) Average coverage of Pol II ChIP-seq, GRO-seq, and 3′ scRNA-seq is plotted for each of the three clusters around true TSSs. Coverages are the median bootstrap estimates of the mean. Browser shots of representative genes from the (B) docked and (C) active clusters are shown. All genes are plotted with their 5′ end to the left regardless of strand.

Docked Pol II Accumulates Upstream of TSSs without Initiating

Metagene analysis reveals that Pol II accumulation is bimodal near true TSSs. Pol II accumulation is evident just downstream of the true TSSs based on both GRO-seq and Pol II ChIP-seq (Figure 3A), as expected given our analysis of scRNA contigs. Surprisingly, and in contrast to the scRNA contigs (Figure 1D), Pol II ChIP-seq coverage around true TSSs is relatively broad, with a mode of the distribution upstream of the TSS (Figure 3A), which is inconsistent with pausing. Examination of individual genes reveals some with Pol II accumulation downstream of the TSS (Figure S3), some with it upstream (Figure S4), and some with it in both places (Figure S2).

Pol II accumulates upstream of TSSs prior to initiation. Because the set of true TSSs we are using was identified from a combination of embryos as well as starved and fed larvae, not all of them are active during L1 arrest. We hypothesized that inclusion of inactive TSSs in metagene analysis resulted in detection of the additional upstream mode of Pol II binding. Consistent with this hypothesis, dividing genes into deciles of scRNA abundance and plotting Pol II ChIP-seq coverage reveal that genes with more scRNA have more Pol II accumulation downstream of the TSS, reflecting paused elongation complexes (Figure 3C). Conversely, genes with less scRNA tend to have accumulation of Pol II upstream of the TSS. A heatmap of individual genes shows this pattern as well (Figure 3D). The same pattern also results when Pol II ChIP-seq coverage is plotted for deciles of GRO-seq signal (Figure S5A) or for deciles of Pol II ChIP-seq coverage in the body of the gene, a proxy for elongation activity (Figure S5B). These results show that accumulation of Pol II upstream of the TSS occurs at genes with relatively little elongation activity. We refer to Pol II accumulated upstream of TSSs as "docked" to indicate that it is recruited to the DNA but has not initiated transcription and that its position is inconsistent with a typical preinitiation complex.

Upstream Accumulation of Pol II Is Not due to Divergent Transcription

We used unsupervised clustering to identify docked genes based on accumulation of Pol II upstream of the TSS. Genes were assigned to one of three clusters based on Pol II ChIP-seq coverage from 200 bp upstream to 200 bp downstream of the TSS. The three clusters include genes with Pol II accumulation upstream of the TSS (docked cluster; 15%), with Pol II downstream of the TSS (active cluster; 14%), and with relatively low amounts of Pol II binding (no signal cluster; 71%) (Figure 4A). Genes in the docked cluster have relatively little Pol II in the body of the gene and very little elongation activity based on scRNA abundance and GRO-seq. Individual examples of genes from the docked cluster clearly show this pattern as well (Figures 4B and S4). Genes in the active cluster have significantly more Pol II in the body of the gene and much more elongation activity based on scRNA abundance and GRO-seq (Figures 4A and 4C). The active cluster is enriched for genes that appear paused based on scRNA abundance (at least 20 reads) and GRO-seq (pause index greater than two) (Fisher's exact test; $p < 2 \times 10^{-16}$ and $= 3.4 \times 10^{-5}$, respectively; Table S2). Indeed, the active cluster appears paused on average, though not all genes in this cluster show evidence of pausing.

Pol II accumulation upstream of TSSs (docking) is not due to antisense or divergent transcription. We used our GRO-seq data to identify genes that are divergently transcribed during L1 arrest and to split them into groups based on the amount of divergent signal. The frequency of divergent transcription has been analyzed elsewhere (Kruesi et al., 2013; Chen et al., 2013). For genes in the docked and active clusters, the Pol II peak position is unaffected by divergent transcription (Figure 5A). We used a ChIP-seq normalization method that improves peak resolution (Enroth et al., 2012), and the offset between the Pol II peak in each cluster is clear (Figure 5A). We also examined the average Pol II position relative to sense and antisense elongation based on GRO-seq and scRNA-seq. For docked genes, Pol II ChIP-seq signal peaks between regions of active transcription, though for active genes, the peak coincides with sense transcription (Figures 5B and 5C). These results are inconsistent with divergent transcription causing Pol II accumulation

Cell PRESS
Open ACCESS

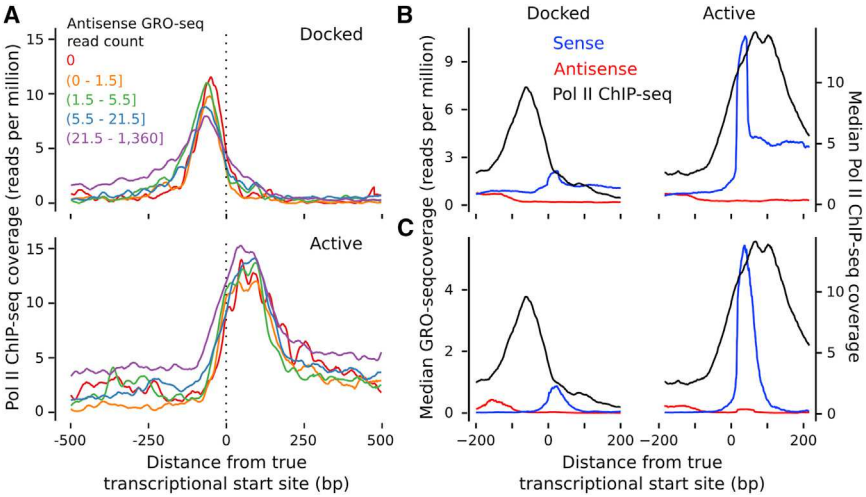

Figure 5. Divergent Transcription Does Not Account for Accumulation of Pol II Upstream of TSSs

(A) Pol II ChIP-seq coverage around 5,192 true TSSs is plotted. Genes are first divided by whether they have antisense GRO-seq reads 200 bp upstream of the TSS, and those that do are grouped by quartiles of antisense read count.

(B and C) Coverage of Pol II ChIP-seq (black), (B) 3′ scRNA-seq, and (C) GRO-seq is plotted around true TSSs of genes in the docked (left) and active (right) cluster. Antisense coverage (red) and sense coverage (blue) are plotted separately.

Coverages for (A)–(C) are the median bootstrap estimates of the mean.

upstream of the TSS. The relative magnitude of Pol II ChIP-seq signal compared to elongation in either direction is also much larger in the docked cluster (Figures 5B and 5C), providing additional evidence that Pol II accumulation upstream of the TSS is not associated with sense or antisense transcription.

Docked and Paused Genes Have Distinct Function and Expression

Genes with docked and active Pol II encode functionally distinct proteins. Genes in the docked cluster are enriched for Gene Ontology (GO) terms associated with growth and development, such as "growth," "larval development," and "translation" (Figure 6A; Table S3). In contrast, genes in the active cluster, many of which are paused, are enriched for GO terms associated with the starvation response, such as "response to stress" and "response to unfolded protein" (a term that includes multiple chaperone proteins). This analysis supports the view that docked and paused genes encode functionally distinct products.

Because docked and active genes are transcribed at different levels during starvation and have different functions, we hypothesized that their transcriptional response to feeding differs. We used previously analyzed microarray data to address this possibility (Baugh et al., 2009). These data were collected from precisely staged animals that hatch in the presence or absence of food, so they either initiate L1 development or enter L1 arrest. We also analyzed expression during recovery from L1 arrest by feeding after 12 hr starvation. Genes docked during starvation are upregulated relative to other genes when development is initiated after hatching with food, consistent with their GO term enrichments (Figure 6B). In contrast, genes in the active cluster are downregulated during development, consistent with them comprising the starvation response. Likewise, docked genes are upregulated and active genes are downregulated at 3 hr of recovery after 12 hr of starvation (Figure 6B). Notably, paused genes in the active cluster show the same pattern of upregulation during starvation and downregulation during recovery (data not shown). This indicates that paused genes do not have a different pattern of expression from the rest of the active cluster. Taken together, these results show that docked and paused genes respond on average in opposite ways to feeding.

mRNA-seq analysis of L1 arrest and recovery confirms that docked and active genes have opposite transcriptional responses to feeding. These data were collected after ~12 hr of L1 arrest (0 hr recovery) and at 1, 3, and 6 hr of recovery by feeding (Maxwell et al., 2012). Docked genes significantly increase transcript abundance during recovery from L1 arrest, and active genes significantly decrease abundance (data not shown), similar to the results of microarray analysis. Yet, average expression does not reveal what fraction of docked and active genes are up- and downregulated, respectively. Therefore, we determined which genes increased or decreased expression during recovery and which of those are statistically significant ($Q < 0.05$). We then intersected these gene sets with the docked and active genes identified by cluster analysis (Figure 6C). Of docked genes, 39% are significantly upregulated during recovery from L1 arrest (Fisher's exact test, $p = 2.4 \times 10^{-7}$), representing 73% of differentially expressed docked genes. Conversely, 45% of active genes are significantly downregulated (Fisher's exact test; $p = 2.2 \times 10^{-7}$), representing 63% of differentially expressed active genes. These results show that the average upregulation and downregulation of docked and active genes, respectively, during recovery from L1 arrest reflect the major trend in each case.

Docking Decreases during Recovery but Is Not Specific to Starvation

Docking decreases in response to feeding. To examine whether docking is restricted to starvation, we analyzed a different Pol II ChIP-seq data set generated from L1 arrest and 1 hr recovery (Baugh et al., 2009). We found a similar number of docked genes during L1 arrest in these data as in the data from Zhong et al. (2010) that we analyze elsewhere (Figure S6A; Supplemental Information). This result shows that docked Pol II is detected in two independent data sets and at a set of common genes (Fishers' exact test, $p < 2.2 \times 10^{-16}$). However, there are eight-times fewer docked genes after 1 hr of feeding. Likewise, the total amount of Pol II upstream of docked genes decreases significantly after 1 hr of feeding compared to L1 arrest (Paired

Cell PRESS
Open ACCESS

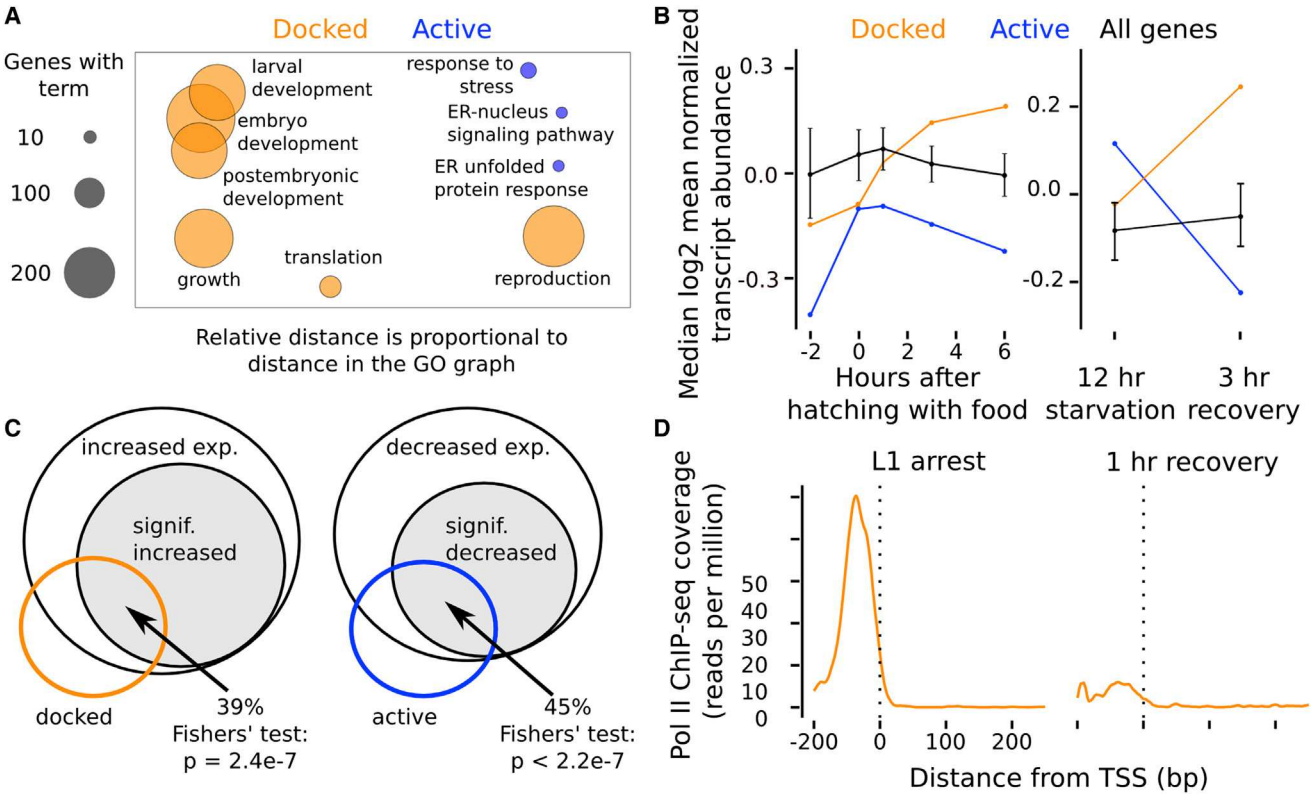

Figure 6. Docked and Active Genes Have Different Functions and Nutrient-Dependent Regulation

(A) Functional enrichments are plotted for the active and docked cluster using the online service "Revigo," which arranges GO terms using multidimensional scaling based on their position in the GO graph. Points are colored by whether they are enriched in the "docked" or the "active" cluster (corrected hypergeometric p value <0.01). The size of the point is scaled according to how many genes are annotated with that functional term in the cluster.

(B) Gene expression during early L1 development (left) and L1 arrest and 3 hr recovery (right) is plotted for the docked (orange) and active (blue) clusters, as well as for all genes (black). Vertical bars on the "All genes" line show the 95% confidence interval of the mean constructed by subsampling 10% of the data 1,000 times.

(C) Venn diagrams showing the numbers of genes in the docked and active clusters whose expression increases, and increases significantly during the first 6 hr of recovery from starvation based on mRNA-seq analysis. Genes were considered "differentially expressed" at a FDR of 1%. We tested for differential expression in 3,093 genes that had detectable mRNA reads (FDR 1%) and also had true TSS calls.

(D) Coverage of Pol II ChIP-seq data around docked genes is plotted during L1 arrest and after 1 hr recovery. In contrast to other figures, Pol II ChIP-seq data are from Baugh et al. (2009). All coverages are the median bootstrap estimates of the mean.

See also Figure S6 and Table S3.

Wilcox test, $p = 9.2 \times 10^{-9}$). This difference is readily apparent in the metagene analysis of docked genes during L1 arrest and 1 hr recovery (Figure 6D). These results show that docking is uncommon during early L1 development compared to L1 arrest, demonstrating that it is influenced by nutrient availability.

Although most genes that are docked during L1 arrest are no longer docked after 1 hr feeding, some genes remain docked (Figure S6A). Interestingly, transcript abundance increases most for these genes during early L1 development (Figure S6B). These genes are also significantly more upregulated after 1 and 6 hr of recovery from L1 arrest (Wilcox test, $p = 0.042$ and $p = 0.001$, respectively). This further illustrates a correlation between docking and upregulation in response to feeding, and it shows that docking can occur outside of starvation.

Independent of starvation, there are other periods in the life cycle associated with lack of or reduced growth. For example, there is no growth during embryogenesis, and the relative growth rate decreases in the latter portion of each larval stage (Knight et al., 2002). Genes associated with growth are not abundantly expressed during embryogenesis or L1 arrest (Zaslaver et al., 2011), and they are downregulated during late L1 development as larvae prepare to molt (Baugh et al., 2009). Likewise, although docked genes are upregulated during early L1 development (0–6 hr; Figure 6B), their expression decreases during later L1 development (6–16 hr; Figure S6B). We examined published Pol II ChIP-seq data prepared from mixed-stage embryos and fed L3 larvae (Gerstein et al., 2010; Zhong et al., 2010). These data were prepared with transgenic animals expressing a GFP fusion to the large Pol II subunit (AMA-1) and an antibody against GFP. This result therefore provides an important control for our results. Docking is nearly as common in embryos as in starved L1s, whereas it is less common in fed L3 larvae (Figure S6C; Supplemental Information). An overlapping set of genes is also docked in each stage (Figure S6D; Fisher's exact test, $p < 2 \times 10^{-16}$ for all pairwise comparisons). It is unclear whether the fed L3 larvae were collected

Please cite this article in press as: Maxwell et al., Pol II Docking and Pausing at Growth and Stress Genes in *C. elegans*, Cell Reports (2014), http://dx.doi.org/10.1016/j.celrep.2014.01.008

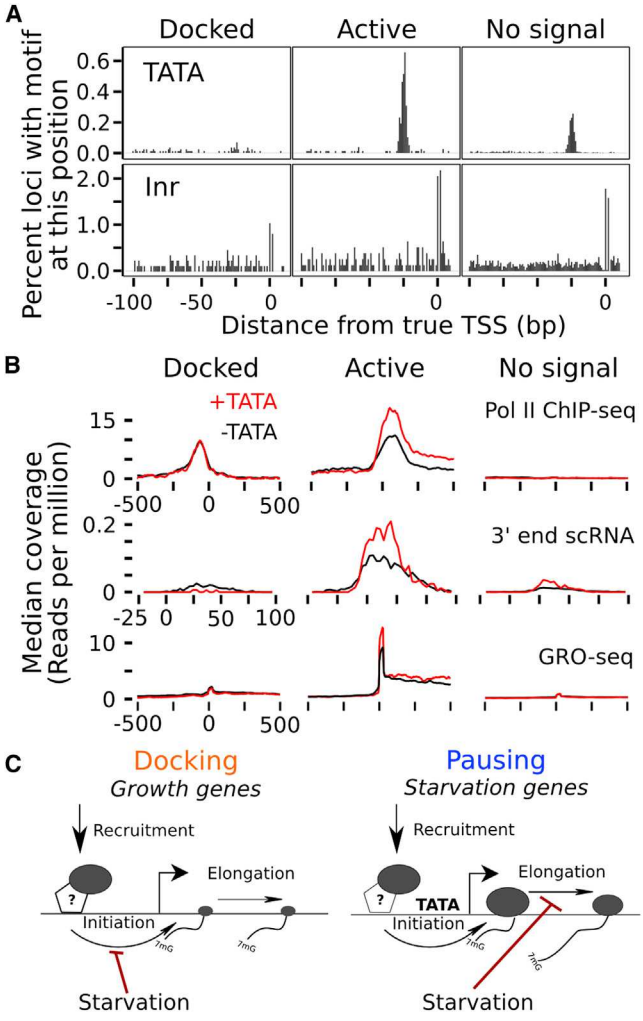

Figure 7. Docked and Active Genes Have Distinct Sets of Core Promoter Motifs

(A) The positional frequency of the Inr and TATA motifs is plotted relative to true TSSs for each of the three clusters.

(B) The coverage of Pol II ChIP-seq, the 3′ ends of scRNAs, and GRO-seq around true TSSs for genes in each cluster is plotted. Genes are split by whether or not they have a canonical TATA motif.

(C) Pol II initiation and elongation are differentially regulated for growth and starvation genes. Upstream accumulation of uninitiated Pol II (docked) is associated with growth and development genes not expressed during starvation but upregulated by feeding. In contrast, promoter-proximal pausing of early elongation is associated with genes expressed during starvation and downregulated during growth, which includes stress-response genes. Starvation genes are much more likely than growth genes to have a TATA box, suggesting alternative core transcriptional machinery in the preinitiation complex of these two sets of genes. We propose that upstream accumulation of docked Pol II involves at least one unknown factor that docks Pol II, represented by a pentagon.

See also Tables S4 and S5.

early or late in the larval stage, and we may detect docking at least in part due to the staging of these animals. These data further show that docking can occur outside of L1 arrest, perhaps in conjunction with reduced growth.

Docked and Paused Genes Have Distinct Promoter Architectures

The fact that docked and active genes have opposite transcriptional responses to feeding suggests a fundamental difference in their regulation. We hypothesized that this difference in regulation is reflected in the promoter architecture of these two gene sets. We used the software FIMO from the MEME suite to look in promoters (200 bp upstream and 100 bp downstream of TSSs) for occurrences of known core motifs defined in the JASPAR database (Bryne et al., 2008; Grant et al., 2011). Most core motifs (9 of 13) are enriched in the active cluster with a FDR cutoff of 5%, including TATA, GC box, and Initiator (Inr) (Fisher's exact test, $Q = 1 \times 10^{-16}$, 4×10^{-10}, and 0.009, respectively; Figure 7A; Table S4). Only MTE-1 is enriched in the docked cluster (Fisher's exact test, $Q = 0.003$), and both Inr and TATA are significantly depleted from this cluster (Fisher's exact test, $Q = 3.4 \times 10^{-6}$ and 5×10^{-8}, respectively). Within the active cluster, paused genes are significantly more likely to have a TATA motif in their promoter than genes that are not paused (45% versus 30%, respectively; Fishers' exact test, $Q = 0.009$). These results show that known core promoter motifs are associated with paused but not docked genes.

Docked genes are associated with distinct promoter motifs. We used the motif identification software DREME from the MEME suite to find motifs differentially enriched among genes in the docked cluster compared to the active cluster, and vice versa (Bailey, 2011). DREME identified 9 motifs enriched in the docked cluster and 11 in the active cluster (Table S5). Consistent with our analysis of known core motifs, this unbiased approach identified the canonical TATA motif TATAWAAG as enriched in the active cluster compared to the docked cluster (DREME; $E = 1.2 \times 10^{-14}$). These results provide additional evidence that genes in the docked and active clusters have distinct sets of core promoter motifs.

The presence of a TATA box has functional consequences at active but not docked genes. The TATA motif is depleted from the docked cluster, but it does occur at some genes in the cluster. However, the presence of TATA does not appear to affect recruitment or elongation at these genes as it does for active genes. That is, TATA is associated with greater Pol II occupancy at active genes based on Pol II ChIP-seq, scRNA-seq, and GRO-seq coverage, but it does not have this effect on docked genes in the rare cases that it is present (Figure 7B). Because active genes do not appear to be regulated at the level of initiation, this result is consistent with TATA promoting recruitment of Pol II to active but not docked genes.

DISCUSSION

We present an integrated genome-wide analysis of Pol II binding, elongation activity, mRNA abundance, and core promoter motifs that reveals distinct forms of postrecruitment regulation of growth and stress genes (Figure 7C). We confirm that promoter-proximal pausing occurs during starvation in *C. elegans*, but it is associated with active stress-response genes that are downregulated in response to feeding. Our results also suggest that initiation is regulated postrecruitment during starvation, resulting in accumulation of docked Pol II upstream of TSSs. In

contrast to paused Pol II, docked Pol II is associated with growth and development genes that are rapidly upregulated in response to feeding. We propose that postrecruitment regulation of initiation and elongation coordinates gene expression with nutrient availability and growth.

Promoter-Proximal Pausing in *C. elegans* L1 Arrest

Like *S. cerevisiae*, the *C. elegans* genome does not encode homologs for any NELF subunits (Narita et al., 2003), an important regulator of pausing in *Drosophila* and mammals (Nechaev et al., 2010). Nevertheless, multiple lines of evidence suggest that promoter-proximal pausing occurs during starvation in *C. elegans*. We show that scRNAs are produced and that their 3′ ends coincide with Pol II accumulation in promoter-proximal regions, consistent with pausing. Pausing has been reported in L1 arrest based on GRO-seq (Kruesi et al., 2013), and we corroborate and expand on this result using scRNA-seq as an independent approach to detect Pol II elongation. scRNA-seq also allows us to show that Pol II typically pauses 30–65 bp downstream of TSSs, similar to *Drosophila* (Nechaev et al., 2010). We also show that the general transcription factor TFIIS has conserved function in *C. elegans*, alleviating backtracking of paused polymerase (Adelman et al., 2005; Kettenberger et al., 2003; Nechaev et al., 2010). Furthermore, the core promoter motifs TATA and Inr are associated with pausing, consistent with *Drosophila* and the complex interaction model that posits a role of these core promoter elements in regulation of early elongation (Kwak et al., 2013). We considered the possibility that our results could reflect premature termination rather than pausing. However, our 3′ scRNA-seq reads (including those that did not map) show no evidence of polyadenylation (data not shown). Given this negative result and the strong similarities to pausing in other systems, we conclude that pausing occurs during L1 arrest in *C. elegans*. Given the apparent lack of NELF, this conclusion implies a NELF-independent pausing mechanism, perhaps involving some combination of the conserved DSIF complex, an unidentified GAGA factor or M1BP homolog, and the core promoter factors (Li and Gilmour, 2013; Missra and Gilmour, 2010; Wada et al., 1998; Kwak et al., 2013).

Pausing is less common in *C. elegans* than in other systems. Our results suggest that ~8% of the 5,192 *C. elegans* genes we examined are paused during L1 arrest. Pausing is even less common in embryos and fed L3 stage larvae (Kruesi et al., 2013). Investigation of *Drosophila* and mammals suggests that pausing is substantially more widespread (Core et al., 2008, 2012; Min et al., 2011; Muse et al., 2007; Zeitlinger et al., 2007). Despite the relatively low frequency of genes with strong evidence for pausing, 5′ accumulation of active Pol II just downstream of TSSs is a pervasive pattern in the data we present. We show that this pattern is not an effect of outliers, suggesting that regulation of elongation may be widespread but with only a modest effect at most genes. It would be valuable to inhibit P-TEFb to determine if in fact many more genes are regulated during early elongation than our studies have revealed (Rahl et al., 2010).

Pausing has been suggested to facilitate a rapid response to stimulus (Lis, 1998; Muse et al., 2007; Zeitlinger et al., 2007), but this is not always the case. Pausing was discovered in the context of the heat shock response, a classic example of rapid stimulus-response dynamics (O'Brien and Lis, 1991; Rougvie and Lis, 1988). However, several studies of stimulus-response systems suggest that paused genes are not necessarily induced in response to stimuli (Gilchrist et al., 2012; Hah et al., 2011; Lin et al., 2011). In mammalian cells, pausing appears to regulate the expression of rapidly induced targets of TNF-α signaling, but not targets of E2 signaling (Danko et al., 2013). These studies and others suggest that the physiological role of pausing depends on context (Adelman and Lis, 2012).

We show that Pol II pauses at actively transcribed genes during starvation in *C. elegans* and that these genes are downregulated in response to feeding. Genes with the most elongation activity, as assessed by scRNA-seq, GRO-seq, and Pol II binding to the gene body, show a pervasive pattern of Pol II accumulation indicative of pausing. This suggests that pausing does not repress transcription during starvation. Furthermore, rather than providing a mechanism to anticipate future activation, genes associated with pausing are downregulated relative to other genes during recovery from starvation. Consistent with this expression pattern, genes associated with pausing are enriched for stress-response genes. Pausing is much less common in embryos or fed larvae (Kruesi et al., 2013), consistent with it reflecting a stress response. These observations suggest that the physiological function of pausing in *C. elegans* is to promote the expression of genes needed during starvation, not to prime genes for induction in response to feeding.

Docking Represents Postrecruitment Regulation of Initiation

Surprisingly, inactive Pol II associates with DNA upstream of TSSs. The amount of upstream Pol II is inversely proportional to the elongation activity at that gene as measured by scRNA-seq, GRO-seq, or Pol II binding to the gene body. We hypothesize that Pol II accumulation upstream of these genes represents nutrient-dependent regulation of transcription initiation. We found the canonical TATA motif in *C. elegans* ~30 bp upstream of the TSS, but Pol II accumulates further upstream (~60 bp). It is unlikely that this accumulation corresponds to a fully assembled preinitiation complex. However, partially assembled preinitiation complexes have been reported by Esnault et al. (2008). We suggest that "docking" should be used to specify recruitment of Pol II upstream of TSSs without initiation.

Three alternative hypotheses could explain upstream accumulation of Pol II: (1) antisense or divergent transcription, (2) abortive initiation, and (3) unregulated transient interaction of Pol II with DNA. The position of docked Pol II is unaffected by and inconsistent with divergent transcription. In addition, the position of docked Pol II upstream of the TSS suggests that it is not undergoing abortive initiation. Multiple lines of evidence argue against a transient interaction model. In this model, Pol II is transiently associating with the relatively weak core promoters of "TATA-less" genes during starvation, with insufficient ATP to promote initiation. However, significant amounts of Pol II accumulate at docked genes compared to paused genes, which is inconsistent with transient interaction. Furthermore, not all genes actively transcribed during L1 arrest have a TATA element in their promoters. This observation shows that TATA is not required for

transcription during starvation, suggesting that relatively weak core promoters can initiate despite starvation. Finally, docking is not confined to starvation because it is present in both embryos and fed L3 larvae. Taken together, our results suggest that upstream accumulation of Pol II is due to it stably associating with DNA but not initiating transcription.

Other examples of postrecruitment, preinitiation regulation of Pol II have been reported. Notably, lymphocyte activation, which involves transcriptome amplification, was recently shown to involve widespread, TFIIH-dependent promoter melting (Kouzine et al., 2013). Pol II also accumulates upstream of many *S. cerevisiae* genes during stationary phase, where it anticipates future induction upon addition of fresh media (Radonjic et al., 2005). Pol II colocalizes with mediator subunits upstream of the TSS during stationary phase, which may provide a binding platform for docked Pol II during quiescence (Andrau et al., 2006). It is tempting to speculate that a similar mechanism operates in *C. elegans*, presumably also involving TFIIH and promoter melting.

Docking Is Associated with Genes Induced by Feeding

We previously reported 5′ accumulation of Pol II at growth and development genes during starvation in *C. elegans*, but we could not distinguish between docking and pausing (Baugh et al., 2009). Pausing was the only form of promoter-proximal postrecruitment regulation known in metazoans, and we speculated that genes with 5′ accumulation of Pol II were paused. However, most of the genes identified were actually docked. Here, we used unsupervised clustering to identify genes with docked Pol II during L1 arrest, and we found that about 15% of genes have docked Pol II. Although docked genes have little transcriptional activity during starvation, a significant fraction of Pol II near TSSs is docked compared to paused, suggesting physiological significance. GO term enrichments suggest that genes with docked Pol II function in growth and development. Consistent with this interpretation, these genes are upregulated in response to feeding after hatching with food and during recovery from L1 arrest. Our results show that Pol II accumulation upstream or downstream of TSSs marks two very different sets of genes.

Docking is influenced by nutrient availability but occurs outside of L1 arrest. Based on the available data, docking is most common in L1 arrest and embryos, and growth gene expression is relatively low in both stages (Zaslaver et al., 2011). Docking also decreases dramatically during immediate recovery from L1 arrest in conjunction with upregulation of docked genes. We propose that docking is associated with periods of no growth (e.g., embryogenesis and L1 arrest) or relatively reduced growth (e.g., the latter portion of each larval stage), but not growth-intensive periods (e.g., the beginning of each larval stage). Such association suggests that docking plays a pervasive role in regulating growth gene expression. We speculate that docking maintains an open chromatin state permissive to regulation, analogous to what has been proposed for pausing (Gilchrist et al., 2008, 2010).

Fundamentally Distinct Regulation of Growth and Stress Genes

Stress resistance and growth reflect distinct, often exclusive priorities. There is a clear distinction in *S. cerevisiae* between the genes expressed during stress and growth (Gasch et al., 2000). Furthermore, in *S. cerevisiae*, stress-response genes have canonical TATA motifs and are regulated by SAGA, whereas housekeeping genes tend to be "TATA-less" and regulated by TFIID (Huisinga and Pugh, 2004). Interestingly, TATA-containing genes tend to have a focused transcriptional initiation pattern in *C. elegans* (Chen et al., 2013), suggesting that this difference is conserved. Starvation causes developmental arrest and confers stress resistance in *C. elegans* (Baugh, 2013), and distinct sets of genes are expressed during arrest and development (Baugh et al., 2009; Maxwell et al., 2012). Our results suggest that these two very different types of genes are transcriptionally regulated by distinct postrecruitment mechanisms and that the mode of regulation is correlated with differences in core promoter architecture. In particular, paused genes are enriched for the TATA motif, and docked genes are depleted for this core motif and others. More experiments are needed to determine whether all genes are prone to docking, as demonstrated for pausing in mammals (Rahl et al., 2010). However, we speculate that growth and stress genes in *C. elegans* tend to employ alternative pathways of preinitiation complex formation, perhaps differentially utilizing TFIID and SAGA, affecting their point of postrecruitment regulation and expression.

Growth rate and stress resistance must be balanced to ensure survival and optimize fitness. Mechanisms that control gene expression in response to fluctuating environmental conditions are critical to environmental adaptation, and their disruption can cause cancer, diabetes, and other diseases. Our results suggest that postrecruitment regulation of initiation and elongation affects growth and stress genes reciprocally to coordinate gene expression with nutrient availability and growth. We anticipate that the putative regulatory mechanisms we describe will be conserved with implications for environmental adaptation as well as human health and disease.

EXPERIMENTAL PROCEDURES

Nematodes were cultured and arrested as previously described (Baugh et al., 2009). scRNA-seq libraries were prepared by size selecting total RNA between 30 and 100 nt, treating sequentially with RNA 5′ Polyphosphatase (Epicenter), Terminator 5′-Phosphate Dependent Exonuclease (Epicenter), and Tobacco Acid Pyrophosphatase (Epicenter), and then following the SOLiD RNA-seq protocol (Applied Biosystems) with appropriate modifications to accommodate irregular insert size. Reads were mapped to the *C. elegans* genome (WS210) in color space using Bowtie v.0.12.7 (Langmead et al., 2009). Additional information on analysis procedures can be found in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The NCBI Gene Expression Omnibus accession number for the scRNA-seq data reported in this paper is GSE40161. Accession numbers of other data sets analyzed are in Table S1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Discussion, Supplemental Experimental Procedures, six figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.01.008.

Cell PRESS
Open ACCESS

ACKNOWLEDGMENTS

We thank Sergei Nechaev for providing protocols and advice for scRNA-seq. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). The Ellison Medical Foundation and the National Science Foundation (IOS-1120206) supported this work.

Received: May 20, 2013
Revised: November 1, 2013
Accepted: January 6, 2014
Published: January 30, 2014

REFERENCES

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

Adelman, K., Marr, M.T., Werner, J., Saunders, A., Ni, Z., Andrulis, E.D., and Lis, J.T. (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. Mol. Cell *17*, 103–112.

Allen, M.A., Hillier, L.W., Waterston, R.H., and Blumenthal, T. (2011). A global analysis of *C. elegans* trans-splicing. Genome Res. *21*, 255–264.

Andrau, J.C., van de Pasch, L., Lijnzaad, P., Bijma, T., Koerkamp, M.G., van de Peppel, J., Werner, M., and Holstege, F.C. (2006). Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. Mol. Cell *22*, 179–192.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653–1659.

Baugh, L.R. (2013). To grow or not to grow: nutritional control of development during *Caenorhabditis elegans* L1 arrest. Genetics *194*, 539–555.

Baugh, L.R., Demodena, J., and Sternberg, P.W. (2009). RNA Pol II accumulates at promoters of growth genes during developmental arrest. Science *324*, 92–94.

Bryne, J.C., Valen, E., Tang, M.H., Marstrand, T., Winther, O., da Piedade, I., Krogh, A., Lenhard, B., and Sandelin, A. (2008). JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic Acids Res. *36* (Database issue), D102–D106.

Chen, R.A., Down, T.A., Stempor, P., Chen, Q.B., Egelhofer, T.A., Hillier, L.W., Jeffers, T.E., and Ahringer, J. (2013). The landscape of RNA polymerase II transcription initiation in *C. elegans* reveals promoter and enhancer architectures. Genome Res. *23*, 1339–1347.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Core, L.J., Waterfall, J.J., Gilchrist, D.A., Fargo, D.C., Kwak, H., Adelman, K., and Lis, J.T. (2012). Defining the status of RNA polymerase at promoters. Cell Rep *2*, 1025–1035.

Danko, C.G., Hah, N., Luo, X., Martins, A.L., Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol. Cell *50*, 212–222.

Enroth, S., Andersson, C.R., Andersson, R., Wadelius, C., Gustafsson, M.G., and Komorowski, J. (2012). A strand specific high resolution normalization method for chip-sequencing data employing multiple experimental control measurements. Algorithms Mol. Biol. *7*, 2.

Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van Berkum, N., Boschiero, C., Holstege, F., and Werner, M. (2008). Mediator-dependent recruitment of TFIIH modules in preinitiation complex. Mol. Cell *31*, 337–346.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Mol. Biol. Cell *11*, 4241–4257.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., et al.; modENCODE Consortium (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. Science *330*, 1775–1787.

Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K.B., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev. *22*, 1921–1933.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Gilchrist, D.A., Fromm, G., dos Santos, G., Pham, L.N., McDaniel, I.E., Burkholder, A., Fargo, D.C., and Adelman, K. (2012). Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. Genes Dev. *26*, 933–944.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Hah, N., Danko, C.G., Core, L., Waterfall, J.J., Siepel, A., Lis, J.T., and Kraus, W.L. (2011). A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell *145*, 622–634.

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. Mol. Cell *13*, 573–585.

Kettenberger, H., Armache, K.J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347–357.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. Nature *436*, 876–880.

Knight, C.G., Patel, M.N., Azevedo, R.B.R., and Leroi, A.M. (2002). A novel mode of ecdysozoan growth in *Caenorhabditis elegans*. Evol. Dev. *4*, 16–27.

Kouzine, F., Wojtowicz, D., Yamane, A., Resch, W., Kieffer-Kwon, K.-R., Bandle, R., Nelson, S., Nakahashi, H., Awasthi, P., Feigenbaum, L., et al. (2013). Global regulation of promoter melting in naive lymphocytes. Cell *153*, 988–999.

Kruesi, W.S., Core, L.J., Waters, C.T., Lis, J.T., and Meyer, B.J. (2013). Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. eLife *2*, e00808.

Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950–953.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Li, J., and Gilmour, D.S. (2013). Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. EMBO J. *32*, 1829–1841.

Lin, C., Garrett, A.S., De Kumar, B., Smith, E.R., Gogol, M., Seidel, C., Krumlauf, R., and Shilatifard, A. (2011). Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). Genes Dev. *25*, 1486–1498.

Lis, J. (1998). Promoter-associated pausing in promoter architecture and post-initiation transcriptional regulation. Cold Spring Harb. Symp. Quant. Biol. *63*, 347–356.

Maxwell, C.S., Antoshechkin, I., Kurhanewicz, N., Belsky, J.A., and Baugh, L.R. (2012). Nutritional control of mRNA isoform expression during developmental arrest and recovery in *C. elegans*. Genome Res. *22*, 1920–1929.

Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes Dev. *25*, 742–754.

Missra, A., and Gilmour, D.S. (2010). Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the *Drosophila* RNA polymerase II transcription elongation complex. Proc. Natl. Acad. Sci. USA *107*, 11301–11306.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511.

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D.K., Hasegawa, J., Omori, M., Inukai, N., et al. (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Mol. Cell. Biol. *23*, 1863–1873.

Nechaev, S., and Adelman, K. (2011). Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. Biochim. Biophys. Acta *1809*, 34–45.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. Science *327*, 335–338.

O'Brien, T., and Lis, J.T. (1991). RNA polymerase II pauses at the 5′ end of the transcriptionally induced *Drosophila* hsp70 gene. Mol. Cell. Biol. *11*, 5285–5290.

Radonjic, M., Andrau, J.C., Lijnzaad, P., Kemmeren, P., Kockelkorn, T.T., van Leenen, D., van Berkum, N.L., and Holstege, F.C. (2005). 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.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432–445.

Rasmussen, E.B., and Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. Proc. Natl. Acad. Sci. USA *90*, 7923–7927.

Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H., and Price, D.H. (2001). A highly purified RNA polymerase II elongation control system. J. Biol. Chem. *276*, 42601–42609.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell *54*, 795–804.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F., et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. *12*, 343–356.

Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in *Drosophila*. Genes Dev. *17*, 1402–1414.

Zaslaver, A., Baugh, L.R., and Sternberg, P.W. (2011). Metazoan operons accelerate recovery from growth-arrested states. Cell *145*, 981–992.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.-W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. Nat. Genet. *39*, 1512–1516.

Zhong, M., Niu, W., Lu, Z.J., Sarov, M., Murray, J.I., Janette, J., Raha, D., Sheaffer, K.L., Lam, H.Y.K., Preston, E., et al. (2010). Genome-wide identification of binding sites defines distinct functions for *Caenorhabditis elegans* PHA-4/FOXA in development and environmental response. PLoS Genet. *6*, e1000848.