Guo et al. Genome Biology 2013, 14:133
http://genomebiology.com/2013/14/9/133

Genome Biology

RESEARCH HIGHLIGHT

Runaway transcription

Jiannan Guo, Tiandao Li and David H Price*

Please see related Research article, http://genomebiology.com/2013/14/9/R98

Abstract

A newly demonstrated defect in RNA polymerase II termination caused by *7SK* snRNA knockdown may have revealed a novel mechanism uncoupling RNA processing from transcription.

Transcriptional elongation by RNA polymerase II (Pol II) is promoted by a switch of the positive transcription elongation factor (P-TEFb) from its inactive to its active form, itself a result of release of P-TEFb from a small nuclear ribonucleic particle (snRNP) in which the small nuclear RNA (snRNA) *7SK* is complexed to a number of proteins [1]. In this issue of *Genome Biology*, Castelo-Branco *et al.* examine the effects of knocking down *7SK* in mouse embryonic stem cells, finding an increase in transcription far downstream of normal termination sites [2].

P-TEFb releases the brakes on paused Pol II

Extensive control of eukaryotic gene expression is exerted at the elongation phase of transcription by Pol II [1]. Control of Pol II elongation is characterized by the default action of negative elongation factors, which include 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole sensitivity inducing factor (DSIF), negative elongation factor (NELF), Gdown1, Gdown1 negative accessory factor (GNAF) and transcription factor IIS (TFIIS). The combined action of these factors generates promoter-proximal paused Pol II, which is found engaged in transcription - but held within the first 100 bp - of approximately one-half of mammalian genes. The transition into productive elongation requires the kinase activity of P-TEFb, which causes phosphorylation of DSIF and the loss of NELF. A new set of factors subsequently become associated with Pol II, which then displays a high elongation rate. At the same time, the site of phosphorylation of the carboxy-terminal domain of the large subunit of Pol II changes from predominately Ser5 to Ser2. Productive elongation complexes facilitate the efficient processing of nascent transcripts into mature mRNAs. Finally, once Pol II passes the polyadenylation site at the 3' end of the transcribed gene, it slows and then terminates, and Pol II and the polyadenylated mRNA are then released.

Because of the prevalence of promoter-proximal paused Pol II and the ability of P-TEFb to cause the transition into productive elongation, metazoans have evolved a unique mechanism for the control of P-TEFb [1]. In rapidly growing cells, most of the P-TEFb population is held in an inactive state by an interaction with hexamethylene-bis-acetamide (HEXIM) inducible proteins in the *7SK* snRNP. Active P-TEFb is released from the snRNP when, and likely where, it is needed and can be returned to the snRNP as genes are shut down [3]. By means of an unknown mechanism, P-TEFb is globally released by actinomycin D, ultraviolet light, P-TEFb inhibitors or any treatment that inhibits Pol II elongation. This sudden release leads to a transient increase in transcription of many genes that were previously occupied with paused Pol II. Global release of P-TEFb can also be triggered by knockdown of the *7SK* snRNA, as used by Castelo-Branco *et al.*, or of factors that stabilize the *7SK* snRNA [4].

Knockdown of *7SK* in embryonic stem cells causes defects in termination

To study the global effects of *7SK* repression in mouse embryonic stem (ES) cells, Castelo-Branco *et al.* [2] collected very deep, strand-specific, RNA-seq datasets 6 hours after knockdown using two different small interfering RNAs against *7SK*. Analysis of the data revealed an unusual pattern of RNA expression caused by knockdown of *7SK*; however, because the pattern of RNA signals is dominated by exons found in stable mRNAs, the global differences between the two conditions were subtle. This can be seen in the UCSC Genome Browser tracks of the forward reads before and after knockdown displayed in the top two rows of Figure 1a (compare 'Control for' with '*7SK* KD for'; the data displayed are for a small interfering RNA targeting the 3' region of *7SK*).

* Correspondence: david-price@uiowa.edu
Biochemistry Department, University of Iowa, Iowa City, IA 52242, USA

BioMed Central © 2013 BioMed Central Ltd.

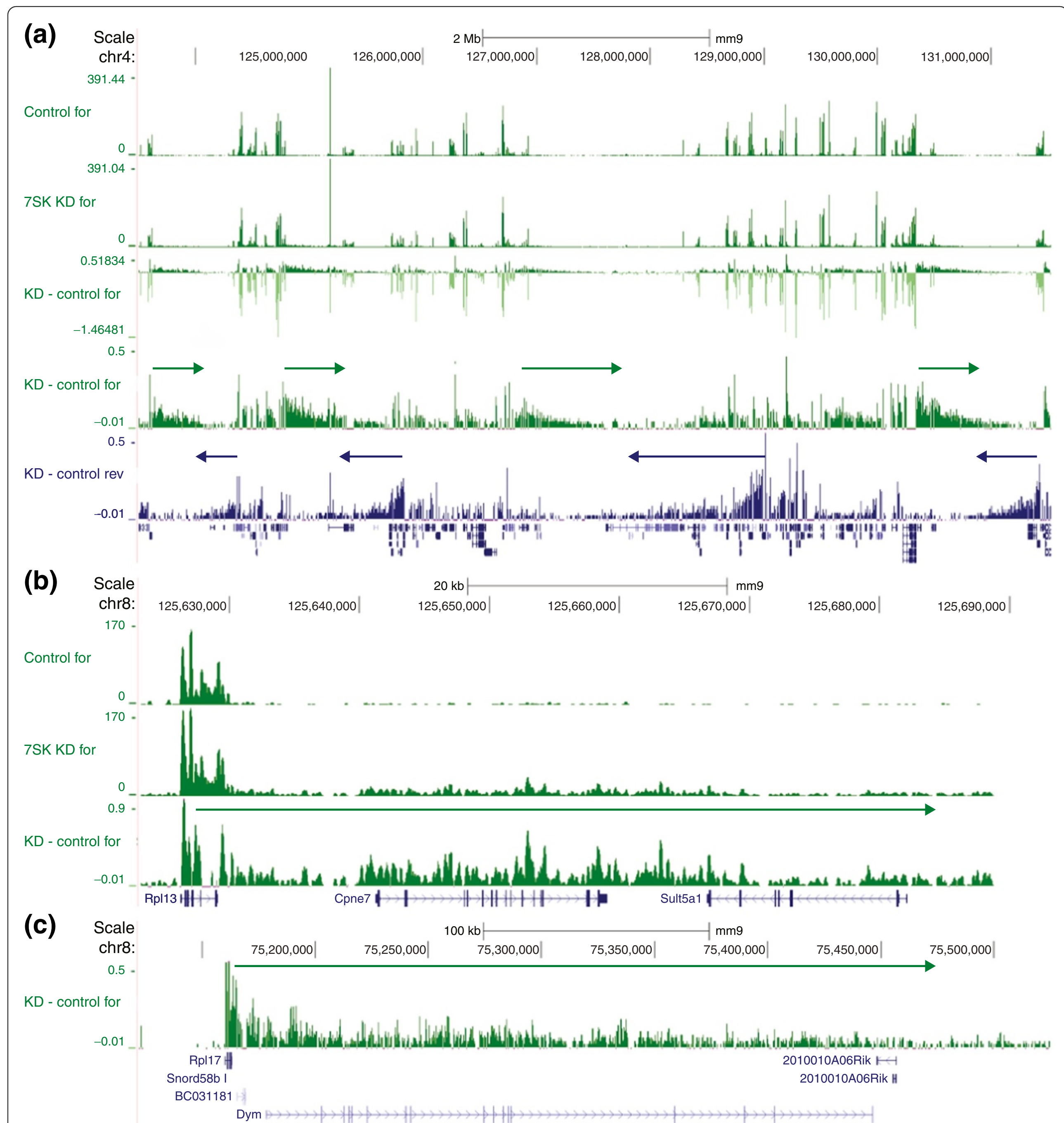

Figure 1 Comparison of RNA-Seq datasets from mouse embryonic stem cells before and after knockdown of *7SK* RNA. (a) UCSC Genome Browser tracks of a representative 6 million bp region showing the original control and *7SK* knockdown RNA-Seq datasets from Castelo-Branco *et al.* [2]. The top two tracks show only the forward direction, with transcription proceeding toward the right. As described in the text, these two tracks were fed into a bioinformatics pipeline that included a subtraction of the control data from the *7SK* knockdown data; the difference track is displayed as the third track (KD - control for), and then again as the fourth track (KD - control for), but in this case with an adjustment to cut-off values below −0.01. A difference track with a −0.01 cut-off was also generated from the reverse reads (KD - control rev, bottom track). Arrows indicate regions of runaway transcription. All tracks in the figure, together with other related datasets, have been deposited in the Gene Expression Omnibus (GSE50860). **(b)** UCSC Genome Browser tracks of a multi-gene region, including a gene encoding a ribosomal protein. The top two tracks display the original data for control and *7SK* knockdown cells (Control for and *7SK* KD for, respectively). The bottom row displays the difference track with a −0.01 cut-off. **(c)** Same display as for **(b)**, but for the multi-gene region surrounding a different ribosomal protein gene. Only the difference track is shown.

To demonstrate the effect of *7SK* knockdown more clearly, we also include in Figure 1 tracks generated from Castelo-Branco *et al.* datasets following a bioinformatics processing pipeline that included normalization and a simple subtraction of the control dataset from the knockdown dataset (Figure 1a, third track).

In the difference track (Figure 1a, third track), the sum of the signals over the entire genome is equal to zero. Positive values indicate transcripts that were induced by *7SK* knockdown. These positive values are mathematically compensated by negative values primarily over exons of pre-existing mRNAs. Negative changes in exon reads should not be thought of as actual decreases in mRNAs. In fact, it is likely that most mRNAs do not change significantly. Therefore, to enhance further the ability to see the *7SK* knockdown-induced transcripts, the difference tracks elsewhere in Figure 1 are presented with a scale that cuts off most of the negative values (Figure 1a, fourth track; Figure 1b, third track; Figure 1c).

A representative 6 million bp region on chromosome 4, with a large number of genes, visually demonstrates the basic finding of the Castelo-Branco *et al.* study (Figure 1a). New transcripts, indicated by arrows, are found spanning several hundred thousand base pairs downstream of expressed genes. Castelo-Branco *et al.* conclude that *7SK* is required for normal transcription termination downstream of polyadenylation sites.

'Trouble ahead, trouble behind'

The difference tracks demonstrate that the new regions transcribed after knockdown of *7SK* do not have discrete stop-points, but instead the density of transcripts tends to trail off after several hundred thousand base pairs. Strikingly, in many of these regions, the same pattern is mirrored by transcription in the opposite direction (Figure 1a, fifth track). We suggest that, perhaps, after knockdown of *7SK*, Pol II enters a 'runaway' transcription mode and, like a runaway train, is doomed to have a head-on collision, derailing transcription in both directions. According to this hypothesis, we would expect long distances to be covered only when significant opposing transcription is not encountered.

Although Castelo-Branco *et al.* suggested that *7SK* possesses a direct role in repressing gene expression as well as transcription downstream of polyadenylation sites, an alternative explanation is that the increase of active P-TEFb caused by the knockdown could be responsible for the effects observed. In support of this idea, Castelo-Branco *et al.* noted that *HEXIM1* expression was upregulated, and this occurs whenever cells are confronted with extra P-TEFb [1]. In this way, P-TEFb would be 'driving that train', as described in the Grateful Dead song commemorating a famous train wreck in 1900 caused by an out-of-control engineer, Casey Jones. Castelo-Branco *et al.* also found 'trouble ahead, trouble behind', in that knockdown of *7SK* led to an increase in long non-coding RNAs (lncRNAs) derived from antisense transcription upstream of some active genes. Such lncRNAs are frequently coordinately regulated with their sense partners [5], and their transcription has been shown to require P-TEFb [6].

Transcription uncoupled from RNA processing

Normally, RNA processing is coupled to transcription, with introns being removed as the gene is being transcribed and 3'-end processing being coupled to termination [7,8]. One of the most striking results extracted from Castelo-Branco *et al.* data is that, after *7SK* knockdown, Pol II appears to ignore polyadenylation signals and does not terminate downstream of active genes. For example, transcription downstream of two ribosomal protein genes is increased after *7SK* knockdown, and the runaway transcription proceeds through downstream genes, again ignoring their polyadenylation signals (Figure 1b,c).

It appears that, after *7SK* knockdown, splicing is also disallowed as the accumulated RNAs in the downstream genes are uniform throughout both introns and exons (Figure 1b,c). In a related study, disruption of the *7SK* snRNP in zebrafish has been demonstrated to affect alternative splicing [9].

Looking forward

The results described by Castelo-Branco *et al.*, together with our analysis of their data, suggest that knockdown of *7SK* causes runaway transcription that is disconnected from RNA processing. This surprising finding provides a window through which the mechanisms of coupling RNA processing to transcription and the consequences of uncoupling might be viewed. The runaway transcription uncovered by Castelo-Branco *et al.* could also be the explanation for pervasive transcription in which large parts of mammalian genomes are transcribed [10].

Currently, it is unclear how knockdown of *7SK* and the resulting increased levels of P-TEFb caused the observed phenotype, but there are several possibilities. If the normal mechanisms regulating recruitment of P-TEFb were bypassed by global release of the factor from the *7SK* snRNP, as a result of the knockdown, the resulting elongation complexes or the transcripts themselves might not be appropriately modified. It will be interesting to determine whether the runaway elongation complexes acquire carboxy-terminal domain phosphorylation at Ser2, which is associated with productive elongation. A second intriguing possibility is that 5' capping of the transcript might not be complete. If transcripts were uncapped, they would likely be unstable. However, if they were capped, but not methylated on the

7 position (m^7G), they would be stable, but unable to be processed efficiently.

If P-TEFb is the engineer 'driving that train', then the *7SK* snRNP should be considered to be the conductor whose responsibility it is to ensure that the train leaves the station at the right time. This new study suggests that uncontrolled P-TEFb might lead to runaway transcription and that the warning from the *7SK* snRNP is: 'Casey Jones, you'd better watch your speed!'

Abbreviations

Bp: Base pair; DSIF: 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole; HEXIM: Hexamethylene-*bis*-acetamide inducible protein; lncRNA: Long non-coding RNA; NELF: Negative elongation factor; Pol II: RNA polymerase II; P-TEFb: Positive transcription elongation factor; Ser: Serine; snRNA: Small ribonucleic acid; snRNP: Small nucleoribonucleic particle; UCSC: University of California, Santa Cruz.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by NIH grant GM35500 to DHP and American Heart Association Fellowship 12POST12040106 to JG.

Published: 30 September 2013

References

1. Guo J, Price DH: **RNA Polymerase II Transcription Elongation Control.** *Chem Rev* 2013, doi:10.1021/cr400105n.
2. Castelo-Branco G, Amaral P, Engström PG, Robson SC, Marques SC, Bertone P, Kouzarides T: **The non-coding snRNA *7SK* controls transcriptional termination, poising and bidirectionality in embryonic stem cells.** *Genome Biol* 2013, **14**:R98.
3. Prasanth KV, Camiolo M, Chan G, Tripathi V, Denis L, Nakamura T, Hubner MR, Spector DL: **Nuclear organization and dynamics of *7SK* RNA in regulating gene expression.** *Mol Biol Cell* 2010, **21**:4184–4196.
4. Krueger BJ, Jeronimo C, Roy BB, Bouchard A, Barrandon C, Byers SA, Searcey CE, Cooper JJ, Bensaude O, Cohen EA, Coulombe B, Price DH: **LARP7 is a stable component of the *7SK* snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated.** *Nucleic Acids Res* 2008, **36**:2219–2229.
5. Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA: **Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells.** *Proc Natl Acad Sci U S A* 2013, **110**:2876–2881.
6. Flynn RA, Almada AE, Zamudio JR, Sharp PA: **Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome.** *Proc Natl Acad Sci U S A* 2011, **108**:10460–10465.
7. Darnell JE Jr: **Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture.** *RNA* 2013, **19**:443–460.
8. Anamika K, Gyenis A, Tora L: **How to stop: the mysterious links among RNA polymerase II occupancy 3' of genes, mRNA 3' processing and termination.** *Transcription* 2013, **4**:7–12.
9. Barboric M, Lenasi T, Chen H, Johansen EB, Guo S, Peterlin BM: ***7SK* snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development.** *Proc Natl Acad Sci U S A* 2009, **106**:7798–7803.
10. Lee JT: **Epigenetic regulation by long noncoding RNAs.** *Science* 2012, **338**:1435–1439.

doi:10.1186/gb-2013-14-9-133
Cite this article as: Guo *et al.*: **Runaway transcription.** *Genome Biology* 2013 **14**:133.