

of H3K27me3 modifications<sup>10</sup>. Expression of somatic genes such as *Hoxa1* and *Hoxb1* is repressed<sup>11</sup>, whereas pluripotent markers *Sox2*, *Pou5f1* and *Nanog* are re-activated<sup>12</sup>. These crucial epigenetic changes essential to allow PGCs to selectively escape the default somatic pathway are tightly associated with expression of Prdm1. Its evolutionarily conserved PR/SET domain closely resembles those present on numerous histone methyltransferases that mediate gene silencing and re-organize nucleosome structure at target genes. In B cells, Prdm1 has been shown to associate with G9a, the predominant histone H3 methyltransferase<sup>13</sup>, whereas complexes with the arginine methyltransferase Prmt5 selectively regulate epigenetic reprogramming in the germ cell lineage<sup>14</sup>. Studies to date have yet to demonstrate the extent to which Prdm1 protein partnerships may be cell-type specific.

#### A new master regulator for germ cells

Saitou and colleagues have exploited single-cell transcriptional profiling strategies to identify a close family member, Prdm14, selectively coexpressed with Prdm1 in the few posterior embryonic cells fated to become

PGCs<sup>2</sup>. As for Prdm1, induction of Prdm14 is dependent on Bmp and Smad signals. Prdm14 function is only required in the germ cell lineage, as homozygous mutant embryos develop normally and adults show no overt tissue abnormalities. However, Prdm14-deficient male and female mice both lack germ cells and are sterile. Loss of Prdm14 results in a failure to upregulate *Sox2* expression, and the mutant PGCs show defects in genome-wide epigenetic reprogramming associated with increased GLP expression and shifted ratios of H3K9me2 and H3K27me3.

It seems likely that additional PRDM family members may have evolved in parallel with the increased complexity of the genome in higher organisms. Silencing default somatic gene expression within the germ cell lineage may simply require more of the same histone modifications. Notably, Prdm14 has been shown to regulate gene expression patterns in human ES cell cultures<sup>15</sup>. The structure of Prdm14 is relatively compact in comparison with that of Prdm1; the distance between the N-terminal PR/SET domain and the C-terminal DNA-binding zinc fingers is roughly half the length. A specific function has yet to be assigned to its proline-rich region, but this distinguishing

feature of Prdm1 may potentially contribute to its broader range of activities in many different cell types. It will be interesting to learn more about discrete and/or possibly overlapping activities shared by Prdm1 and its close family member Prdm14, such as DNA binding-site specificities and selection of partners responsible for reprogramming the transcriptional machinery. We await further studies needed to shed light on the possibly conserved roles played by Prdm1 and Prdm14 in the developing human embryo.

1. Hayashi, K. *et al. Science* **316**, 394–396 (2007).
2. Yamaji, M., *et al. Nat. Genet.* **40**, 1016–1022 (2008).
3. Martins, G. & Calame, K. *Annu. Rev. Immunol.* **26**, 133–169 (2008).
4. Horsley, V. *et al. Cell* **126**, 597–609 (2006).
5. Magnúsdóttir, E. *et al. Proc. Natl. Acad. Sci. USA* **104**, 14988–14993 (2007).
6. Vincent, S.D. *et al. Development* **132**, 1315–1325 (2005).
7. Ohinata, Y. *et al. Nature* **436**, 207–213 (2005).
8. Robertson, E.J. *et al. Development* **134**, 4335–4345 (2007).
9. Surani, M.A. *et al. Cell* **128**, 747–762 (2007).
10. Seki, Y. *et al. Development* **134**, 2627–2638 (2007).
11. Saitou, M. *et al. Nature* **418**, 293–300 (2002).
12. Yabuta, Y. *et al. Biol. Reprod.* **75**, 705–716 (2006).
13. Gyory, I. *et al. Nat. Immunol.* **5**, 299–308 (2004).
14. Ancelin, K. *et al. Nat. Cell Biol.* **8**, 623–630 (2006).
15. Tsuneyoshi, N. *et al. Biochem. Biophys. Res. Com.* **367**, 899–905.

## Mapping the strand-specific transcriptome of fission yeast

Thomas R Gingeras

**Pervasive genome-wide transcription is widespread in eukaryotic cells, but key features of the transcriptome have yet to be fully characterized. A new study using antibody-based detection of RNA-DNA duplexes on tiling arrays now reveals a complex, strand-specific transcriptional world in fission yeast.**

The biochemical evidence for pervasive genome-wide transcription has been well established for many organisms<sup>1,2</sup>. These and many other studies point to a transcriptional organization for many genomes that can be characterized as highly interleaved (Fig. 1)<sup>3</sup>. However, at present, it is still unclear what the biological roles of these previously unannotated transcripts are. One step toward designing useful experiments is to investigate whether the unannotated transcribed regions share

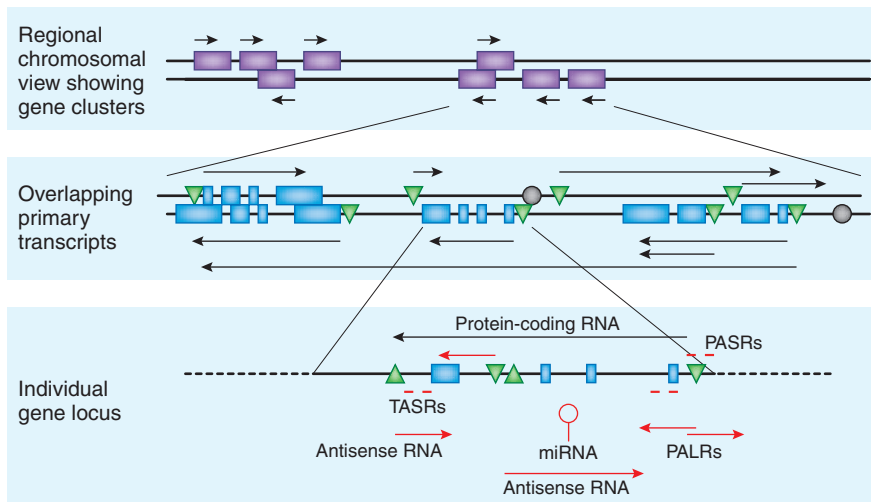
characteristics that would allow for hypotheses to be formulated and genetically tested. Dutrow *et al.*, reporting on page 977 of this issue<sup>4</sup>, provide such a dataset in their analysis of the fission yeast *Schizosaccharomyces pombe* transcriptome.

#### Detecting duplexes

Consistent with data obtained from studies of human<sup>5</sup>, mouse<sup>6</sup> and fly<sup>7</sup>, the authors find that virtually all of the euchromatic genome is transcribed. However, two things of note distinguish this observation from other genome-wide transcription studies. The first involves the technical approach used by the authors to obtain their transcriptome maps. Specifically, the authors used an antibody (S9.6) raised

against RNA-DNA duplexes to identify duplex regions formed on DNA probes that were part of whole-genome tiling arrays. This allowed for the strand-specific designation of each of the detected transcribed regions without the use of reverse transcriptase or RNA-DNA amplification techniques, which can have issues associated with the production of only single-stranded products. The results using this technical approach are notable in both their range of detection (7,600-fold) and their specificity (reduction in hybridization signal of 80- and 20,000-fold with one and two base mismatches, respectively). However, given that these results are dependent upon the use of an immunological approach, there is a concern as to whether these results will be reproducibly achieved by

Thomas R. Gingeras is at the Department of Functional Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 11724, USA.  
e-mail: [tom\\_gingeras@affymetrix.com](mailto:tom_gingeras@affymetrix.com)



**Figure 1** Three levels of resolution showing pervasive interleaved transcription. The clustering of genes in the upper portion of the figure cloaks the overlapping transcription of protein-coding and noncoding transcripts observed within and between genic regions, as depicted in the middle panel. These transcripts also point to multiple regulatory regions (triangles and circles) that are positioned within genes and present on opposite strands. The ultimate fate of some of these long transcripts (for example, promoter-associated long RNAs; PALRs) is to provide short RNAs such as microRNAs, promoter-associated short RNAs (PASRs) and termini-associated short RNAs (TASRs).

other laboratories using even slightly different labeling protocols, antibody preparations from the hybridoma cell line or detection array systems. Similar concerns have plagued chromatin immunoprecipitation studies. Nevertheless, the results achieved using this approach provide a fresh alternative to achieve strand-specific RNA maps using tiling arrays.

The second notable aspect of this study focuses on the characterization of previously unannotated transcribed genomic regions. One class of unannotated transcription in *S. pombe* noted by Dutrow *et al.* involves the detection of widespread antisense transcription. The authors indicate that the detection of antisense transcription in *S. pombe* was less prevalent when polyadenylated (poly(A)+) RNA was mapped compared to total RNA. The authors interpret this to be consistent with the possibility of the antisense transcripts being poly(A)<sup>-</sup> or possessing reduced length polyadenylation. Dutrow *et al.* suggest that the cause of this antisense transcription is opportunistic on the basis of the negative correlation with histone H3 occupancy and the positive correlation of coordinated expression of sense-antisense transcription during conditions of gene expression changes (for example, heat shock) and the presence of histone H3 lysine 36 trimethylation (H3K36me3). However, the conclusion that *S. pombe* antisense transcription is in large measure composed of poly(A)<sup>-</sup> RNA implies that such transcripts would likely have a rapid turnover rate and be relatively short-lived. If correct, this observation would

represent a significant difference compared to that observed in mouse and human cells<sup>3,8</sup>, in which most of the detected antisense transcription is detected in poly(A)<sup>+</sup> RNA samples and contains reasonable length polyadenylation, as revealed by cDNA sequencing. Direct empirical determination of the polyadenylation state of most antisense transcripts in *S. pombe* would be straightforward and would not only confirm these observations but also provide possible insights as to why such a marked contrast is observed between fission yeast and higher eukaryotic cells.

#### Opportunistic or deterministic?

Another characteristic of antisense transcripts observed by Dutrow *et al.* involves the previously uncharacterized transcribed regions flanking the tRNAs residing in the *imr* repeats found in the heterochromatic centromeric regions. The regions flanking the tRNA genes seem to be transcribed on the antisense strand relative to the tRNA gene. Such flanking antisense transcription is reported not to be observed at tRNA genes found in euchromatin regions. Again, the authors see this antisense transcription as the result of opportunistic conditions set up by the directed transcription of the tRNA genes. Extending this line of thought, the authors conclude that the genome-wide baseline transcription observed along intergenic regions is also opportunistic and indicative of the chromatin state of these regions.

Although initially attractive, this explanation places the role of transcription of a

large portion of a genome as a passive and baseline condition in the cell. Such a promiscuous role for transcription is troubling for two reasons. First, as indicated by the authors themselves, the RNA detected in their studies reflects a steady state condition in the cells, and thus, these molecules are not likely to be short-lived. This is especially the case when the same transcribed regions are observed at multiple time points during development or in response to external stimuli, as seen in this study. These nontransient RNAs within cells are thus likely to be immediately associated with a diverse collection of RNA-binding proteins. The roles of these proteins are substantial but their abundance in a cell is not. An organizational strategy that uses opportunism on such a global scale greatly increases the requirement for regulatory complexity to discern the products of opportunism from determinism so as to judiciously use the limited RNA binding–protein resources of the cell and sets up conditions for creating transcripts that have the same regulatory signals as transcripts created in a deterministic fashion.

One of the results from the pilot ENCODE studies may point to a less opportunistic reason for the synthesis of such transcripts<sup>1</sup>. An analysis of ENCODE regions aligned for 23 mammals and 5 other vertebrates showed significant enrichment for short islands of conservation within transcripts of unknown function (TUFs), despite absence of detectable conservation when the enrichment was averaged across the whole length of each transcribed region. Thus, long intergenic or antisense transcripts can be made in a directed and regulated fashion in order to provide short functional RNAs (for example, microRNA primary transcripts) or allow for the rapid evolution of sense-antisense transcript pairs (Fig. 1). It seems likely that additional studies will be undertaken involving transcripts originating from unannotated regions to determine whether stable short RNAs are also found mapping to these same regions and which are enriched in evolutionary conserved sequences, as has been observed in human cell lines<sup>3</sup>.

#### COMPETING INTERESTS STATEMENT

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

1. Birney, E. *et al.* *Nature* **447**, 799–816 (2007).
2. Gingeras, T.R. *Genome Res.* **17**, 682–690 (2007).
3. Kapranov, P. *et al.* *Science* **316**, 1484–1488 (2007).
4. Dutrow, N. *et al.* *Nat. Genet.* **40**, 977–986 (2008).
5. Efroni, S. *et al.* *Cell Stem Cell* **2**, 437–447 (2008).
6. Katayama, S. *et al.* *Science* **309**, 1564–1566 (2005).
7. Manak, J.R. *et al.* *Nat. Genet.* **38**, 1151–1158 (2006).
8. Carninci, P. *et al.* *Science* **309**, 1559–1563 (2005).