*Nature Reviews Molecular Cell Biology* | Published online 21 Mar 2017; doi:10.1038/nrm.2017.25

EPIGENETICS

# DNA methylation prevents intragenic transcription

Cytosine methylation at gene promoters by methyltransferases such as DNA methyltransferase 3B (DNMT3B) results in gene repression. DNMT3B also binds in gene bodies by interacting with trimethylated Lys36 of histone H3 (H3K36me3). Loss of H3K36me3 in gene bodies has tumorigenic effects, but the physiological role of intragenic DNA methylation is poorly understood. Neri *et al.* now show that methylation in gene bodies prevents aberrant and potentially deleterious intragenic transcription.

The authors generated *Dnmt3b* knockout (*Dnmt3b*$^{-/-}$) mouse embryonic stem cells. Chromatin immunoprecipitation and sequencing revealed that intragenic DNMT3B binding occurs preferentially at highly expressed genes, and whole-genome bisulphite sequencing revealed a significant decrease in 5-methylcytosine levels at exons and introns in *Dnmt3b*$^{-/-}$ cells. Importantly, when compared with wild-type cells, *Dnmt3b*$^{-/-}$ cells exhibited notably more transcription initiation at intermediate exons (that is, from the second exon onwards) than at the first exon; this ratio was considerably reduced following the introduction of wild-type — but not of catalytically inactive — DNMT3B.


> *Dnmt3b*$^{-/-}$ cells exhibited notably more transcription initiation at intermediate exons

To characterize the intragenic transcription start sites (TSSs), mRNAs were enzymatically decapped, leaving a 5′ monophosphate group that was selectively used for adaptor ligation, thereby enabling global mapping of TSSs at single-base resolution. This revealed a significant increase of transcription initiation within the bodies of highly expressed genes in *Dnmt3b*$^{-/-}$ cells. The intragenic TSSs were characterized by a lack of canonical TSS sequence features and by an enrichment of CpG dinucleotides, and of several CG-rich transcription factor binding motifs, in their vicinity.

SET domain-containing protein 2 (SETD2)-dependent H3K36 trimethylation within gene bodies occurs during transcription elongation and was previously shown to prevent intragenic transcription initiation. Depletion of SETD2 resulted in a loss of DNMT3B intragenic binding in wild-type cells and a significant increase in the number of TSSs within gene bodies, of which more than 40% were in common with the TSSs identified in *Dnmt3b*$^{-/-}$ cells. Furthermore, DNMT3B mutants that cannot bind to H3K36me3 were unable to fully restore DNMT3B intragenic binding and activity in *Dnmt3b*$^{-/-}$ cells, and did not reduce aberrant intragenic transcription.

Most of the transcripts generated from the intragenic TSSs were degraded by the nuclear RNA exosome complex, but almost all of the highly expressed intragenic transcripts were polyadenylated and were detected in the cytosol. The stability of these aberrant transcripts was found to be equivalent to that of RNAs transcribed from canonical TSSs. Importantly, ribosome profiling and active mRNA translation sequencing revealed that *Dnmt3b*$^{-/-}$ cells display a significant reduction in ribosome occupancy at 5′ untranslated regions and a marked increase in ribosome occupancy at the introns of highly expressed genes, indicating that increasing intragenic transcription may generate aberrant proteins and have harmful physiological consequences.

*Eytan Zlotorynski*

**ORIGINAL ARTICLE** Neri, F. *et al.* Intragenic DNA methylation prevents spurious transcription initiation. *Nature* http://dx.doi.org/10.1038/nature21373 (2017)