Cell Formatting Guidelines

Title: No more than 10-12 words

Summary: Single paragraph, no more than 150 words. Should contain the following:

-brief background of the questions (avoid statements about how it’s not well understood)

-A description of results and approaches/model systems

-Indication of the broader significance of the work (Discourage novelty claims)

Introduction: A good introduction is succinct. It should allow the reader to understand the motivation for the study and results. No subheadings.

Results: Divided with specific subheadings

Discussion: Explain the significance of the results and place in a broader context. Should include future directions. Should not be redundant.

DNA Methylation Regulates Alternative Polyadenylation via CTCF and the Cohesin Complex

**Summary (currently 145 words)**

Dysregulation of DNA methylation and mRNA alternative cleavage and polyadenylation (APA) are both prevalent in cancer. We discovered a link between these two processes when we compared genome-wide DNA methylation and polyadenylation site (pAS) usage between methylation-competent HCT116 and methylation-deficient DKO cells. 546 genes harboring more than one pAS undergo APA in association with differential gene body DNA methylation. Removal of DNA methylation enables CTCF binding and recruitment of the cohesin complex and, at a subset of these genes, promotes proximal pAS usage. Importantly, depletion of the RAD21 cohesin complex component, in the DNA demethylated context, recovered distal pAS usage. By querying RNA-seq and DNA methylation data from TCGA, we authenticated the correlation between gene body DNA methylation and pAS isoform expression *in vivo*. Thus, we present a molecular mechanism of DNA methylation-regulated APA and demonstrate how aberrant DNA methylation impacts transcriptome diversity in cancer.

**Introduction**

Normal vertebrate development requires the concerted effort of essential processes such as DNA methylation and alternative cleavage and polyadenylation (APA), both of which contribute to transcriptome diversity. DNA methylation, in which a methyl group is covalently attached to the 5′ position of cytosine, is an ancient and highly conserved mechanism in plants and animals, which exhibit tissue specific methylation patterns. (Jaenisch and Bird, 2003a; Song *et al.*, 2005; Weber *et al.*, 2007; Suzuki and Bird, 2008; Feng *et al.*, 2010; Ghosh *et al.*, 2010; Zemach and Zilberman, 2010; He, Chen and Zhu, 2011; Shenker and Flanagan, 2012a). Knockout of DNA methyltransferases is embryonic lethal in mice (Li, Bestor and Jaenisch, 1992; Okano *et al.*, 1999; Jaenisch and Bird, 2003a). Increased 3′ UTR length positively correlates with more complex multicellular organisms and allows for more nuanced regulation of gene transcription, such as APA. APA can alter the coding regions of genes (intronic or internal exon APA), affect translational efficiency, or affect mRNA stability and localization (3′ UTR or alternative terminal exon APA; Elkon, Ugalde and Agami, 2013). Like DNA methylation, APA is regulated in a cell- and tissue-specific manner and is essential for normal development (Elkon, Ugalde and Agami, 2013; Lianoglou *et al.*, 2013). Despite the necessity of these processes, DNA methylation and APA can be diverted to promote carcinogenesis.

The cancer genome is globally hypomethylated with regional hypermethylation that is often associated with CpG islands (CGI; Jaenisch and Bird, 2003a). Promoter CGI hypermethylation can silence tumor suppressor genes. Although the role of DNA methylation in non-promoter regions is not well understood, it appears to be important, evidenced in part by the global loss of DNA methylation seen in cancer (Baylin and Jones, 2011). We have previously shown that >80% of differentially methylated regions (DMRs) between indolent and aggressive prostate cancers occur at non-promoter regions, such as introns, exons and gene 3′ ends (Bhasin *et al.*, 2015). Misregulation of APA causes both shortening and lengthening of mRNAs in cancer, and different cancer types exhibit different trends in global APA isoform usage (Fu *et al.*, 2011; Morris *et al.*, 2012; Lee *et al.*, 2016). Shortened oncogenes can escape miRNA-mediated degradation (Mayr and Bartel, 2009). Despite the prevalence of APA in cancer, mechanisms regulating isoform usage are not well-established.

Gene body methylation positively correlates with gene expression (Yang *et al.*, 2014). This correlation suggests that genic methylation affects regulation of RNA production, maintenance, and/or processing. Approximately 10% of DMRs are located proximal to gene 3′ ends, where 3′ end maturation, including APA, occurs (Ogorodnikov, Kargapolova and Danckwardt, 2016). We sought to investigate the relationship between DNA methylation and alternative polyadenylation. Poly(A)-seq and MBD-seq in methylation-competent HCT116 and methylation-deficient DKO revealed a significant association between DNA methylation and APA, mediated by binding of CTCF, a methylation-sensitive insulator protein (Bell and Felsenfeld, 2000). Here, we present a mechanism in which DNA methylation regulates aberrant APA in cancer via CTCF and the cohesin complex.

**Results**

*Differential DNA methylation is associated with alternative poly(A) site usage*

To investigate the relationship between differential pAS usage and DNA methylation we performed poly(A)-specific sequencing in two isogenic cell lines: HCT116 and DKO (Figure S1A). HCT116 is a human colorectal tumor cell line. DKO is an HCT116-derivative cell line with DNA methyltransferases 1 and 3b knocked out; consequently, it has less than 5% genomic DNA methylation compared to HCT116 (Rhee *et al.*, 2000, 2002). Comparison of poly(A) replicates showed no global dysregulation of APA , and the frequency of hexamers known to precede pAS is comparable to the poly(A) database (Figure S1B&C; Table S1) (Lee *et al.*, 2007). Based on analysis of quadruplicate poly(A)-seq wemapped 32,245 pAS (13,369 genes) and 25,905 pAS (13,359 genes) in HCT116 and DKO cells respectively (Table 1; Table S2). Among genes using more than one pAS, we identified 546 genes undergoing APA between HCT116 and DKO using criteria that at least one pAS shows relative usage differing by at least 1.5 fold and 10% between the two cell lines (Table 1; Figure 1A; Table S3). RNA-seq confirmed that the observed APA was independent of differential *trans*-acting factor expression (Figure S1D) (Elkon, Ugalde and Agami, 2013; Masamha *et al.*, 2014). Furthermore, we observed that 412 of these APA candidate genes preferentially use the proximal pAS in DKO.

To identify potential mechanisms regulating preferential proximal pAS usage in the absence of DNA methylation, we queried 161 Encode transcription factors for potential binding of the DNA sequences between shifting pASs in the 412 candidate genes (Figure 1B). Of the top ten transcription factors identified, CCCTC-binding factor (CTCF) has well-established methylation-sensitive binding properties (Renda *et al.*, 2007; Phillips and Corces, 2009). Of those genes that could be bound by CTCF, we chose two candidates in which to delineate the mechanism: HEAT Repeating Protein 2 (*HEATR2*) and Nuclear Factor Y subunit A (*NFYA*). *HEATR2* has three pASs, two of which are intronic (Figure 1C). Poly(A)-seq showed a 14.6 fold increase in relative usage of the most proximal, intronic pAS from HCT116 to DKO. HEATR2 mediates dynein pre-assembly to support ciliary motility and mutations in *HEATR2* cause primary ciliary dyskinesia (Horani *et al.*, 2012, 2018; Diggle *et al.*, 2014). *NFYA* has four pASs, all in the 3′UTR (Figure 1D). Poly(A)-seq showed a 2.6 fold decrease in relative usage of the most distal pAS between HCT116 and DKO. NFYA is a subunit of the transcription factor NFY, which recognizes the motif CCAAT (Zambelli and Pavesi, 2017). Many genes overexpressed in cancer cells contain a CCAAT motif bound by NFY, implicating NFY in carcinogenesis (Gurtner, Manni, and Piaggio, 2017). Despite comparable methylation throughout both *HEATR2* and *NFYA*, substantial loss of methylation in DKO was seen specifically at the CGI between our pASs of interest. In both cases, ChIP-seq for CTCF revealed CTCF binding in DKO, but not in HCT116, in alignment with the differentially methylated CGI (Figure 1C&D). This differential binding was specific to the identified CGI; CTCF binding in the remainder of both *HEATR2* and *NFYA* was comparable between the two cell lines. This data suggests an association between DNA methylation and alternative polyadenylation, which we postulate is regulated by CTCF.

*DNA methylation regulates APA via CTCF*

To confirm DNA methylation regulates APA, we treated HCT116 cells with a DNA demethylation agent, 5-aza-2′-deoxycytidine (DAC; Momparler, 2005). Sequencing of bisulfite converted DNA confirmed hypermethylation of HCT116 cells and hypomethylation of DKO cells (Ciark *et al.*, 1994). DAC-treated HCT116 cells exhibited a substantial decrease in fully methylated alleles with 16.7% and 25% remaining in *HEATR2* and *NFYA* respectively (Figure 2A). Isoform-specific qRT-PCR showed that distal isoform production decreased 5.5 fold in *HEATR2* and 1.6 fold in *NFYA* from HCT116 to DKO (Figure 2B). Demethylation of HCT116 cells by DAC treatment showed a 5.8 and 2.5 fold decrease in the distal isoform in *HEATR2* and *NFYA* respectively, mimicking DKO (Figure 2B).

ChIP-qPCR for CTCF detected increased binding in DKO in alignment with the differentially methylated CGI between the proximal and distal pASs in *HEATR2* and *NFYA* (Figure 2C). In both genes, CTCF binds at the same regions in DAC-treated HCT116 cells, confirming that CTCF binds these regions in the absence of DNA methylation. Similarly, we observed increased binding of RNA Polymerase II (Pol2) at the same CGI in *HEATR2* and *NFYA* by ChIP-qPCR (Figure 2C). This data indicates that the absence of DNA methylation allows CTCF to bind, promoting Pol2 buildup and proximal isoform usage. After DAC treatment, HEATR2 protein expression decreased compared to untreated HCT116 (Figure 2D). We postulate that this is due to a shift from the stable, distal isoform to an unstable, proximal isoform. No change was observed for NFYA, indicating that the steady state expression of NFYA protein is not affected. Differential binding of CTCF and Pol2 is a true phenomenon, rather than the result of differential expression (Figure 2D).

To determine whether CTCF binding is necessary and sufficient for methylation-dependent APA, we transfected HCT116 cells with a luciferase reporter construct containing either a wild type *NFYA* 3′UTR (Luc*NFYA*) or a mutant *NFYA* 3′UTR, lacking the CTCF motif (Luc*NFYA*\*; Figure 2E). We expected the Luc*NFYA*\* construct to mimic the methylated *NFYA* allele due to the loss of CTCF binding. Northern blotting against luciferase showed a 1.44 fold increase in expression of the distal isoform in the *LucNFYA*\* transfected cells compared to the Luc*NFYA* transfected cells. The increase in distal APA isoform expression in the absence of the CTCF binding site indicates that CTCF is required for the regulation of DNA methylation-dependent APA. However, this is a smaller shift than observed between HCT116 and DKO cells (Figure 1D; Figure 2B). We postulate that this may be due, in part, to the use of a strong promoter, which preferentially uses proximal pASs (Gruber *et al.*, 2014). More importantly, compounding aspects of chromatin structure that cannot be recapitulated in this assay may be involved in APA regulation.

*CTCF cooperates with the cohesin complex to regulate APA*

CTCF is known to recruit the cohesin complex to assist in the formation of chromatin loops and topologically associating domains (Phillips and Corces, 2009). ChIP-seq for RAD21 cohesin complex component (RAD21) and structural maintenance of chromosomes 1 (SMC1) in DKO cells showed binding of both proteins concomitant with CTCF at the differentially methylated CGI in *HEATR2* (Figure 3A)and *NFYA* (Figure 3B). Minimal binding was observed in methylated HCT116 cells (Figure 3A&B). Surprisingly, binding of RNA Polymerase II phosphorylated at serine 5 of the YSPTSPS repeats (Pol2Ser5) was increased in DKO, but not in HCT116, in the same region of both genes (Figure 3A&B). Pol2Ser5 is responsible for transcription initiation and is usually found at the 5′ end of genes (Bowman and Kelly, 2014). Its presence at the CGI located between the proximal and distal pASs indicates that increased Pol2 occupancy (Figure 2C) is due not only to stalling of the elongating Pol2 phosphorylated at serine 2 (Pol2Ser2) but also the recruitment of a new initiating Pol2Ser5. ChIP-seq for Histone 3 Lysine 27 acetylation (H3K27ac), a promoter and enhancer mark associated with the cohesin complex, showed increased occupancy at the CGI bound by CTCF, RAD21 and SMC1, particularly in *HEATR2* (Figure 3A&B) (Creyghton *et al.*, 2010). These observations suggest that the cohesin complex and Pol2Ser5 are recruited to the unmethylated CGI, stalling transcription and generating the proximal APA isoform. A clustering analysis to identify co-occurring dynamic shifts of CTCF, RAD21, SMC1, Pol2Ser2p and Pol2Ser5p binding, as well as H3K27ac and methylation, identified nine distinct clusters (Figure S2; Tables S4 and S5). *HEATR2* and *NFYA* are contained in cluster 3 (47 genes) and cluster 7 (97 genes) respectively (Figure 3C). Both clusters are characterized by a decrease in DNA methylation in the DKO cell line and concomitant increased binding of CTCF, RAD21, SMC1 and Pol2Ser2p. The increase in CTCF is particularly noticeable in cluster 7 (Figure 3C). Additionally, cluster 3 shows increased presence of H3K27Ac and Pol2Ser5p (Figure 3C).

To directly test the role of the cohesin complex in methylation-dependent APA, we made use of an auxin-inducible RAD21-degron system derived from HCT116 cells. RAD21 alleles in these cells are tagged with a destabilizing domain (degron), causing ubiquitination and degradation after exposure to auxin (3-indoleacetic acid, IAA; Natsume *et al.*, 2016). RAD21 was substantially depleted after treatment with IAA for 8 hours in these cells while DAC treatment resulted in DNA methylation and loss of DNMT1 (Figure 3D). IAA treatment alone does not change isoform production. However, the shift in isoform production was seen in newly transcribed RNA after isolation by 4-thiouridine (4sU) treatment (Rädle *et al.*, 2013; Zheng *et al.*, 2018). A 94% depletion of RAD21 (Figure 3D) in DAC-treated cells recovered distal isoform production 1.78 fold in *HEATR2* and 1.81 fold in *NFYA* compared to demethylated cells with intact RAD21 expression (Figure 3E). We do not observe the same recovery at the protein level, likely due to the brevity of the treatment (Figure 3D). Consistent expression of CTCF between samples confirms that the recovery was exclusively due to loss of RAD21 (Figure 3D). These results confirm that RAD21 is necessary for DNA-methylation regulated APA and are consistent with the cohesin complex being involved in the regulation.

*DNA methylation-regulated APA is evident in vivo in TCGA cohorts*

Having established a connection between DNA methylation and APA *in vitro*, we sought to find evidence of the association *in vivo*. We analyzed 450k Illumina methylation array and RNA-seq data of 11 cancer cohorts from TCGA and a merged set consisting of data points from all 11 cohorts to identify any cohort-specific correlations, as well as correlations across cancer types. Although *NFYA* showed only modest correlation in 2/12 cohorts, *HEATR2* showed a strong correlation between DNA methylation and APA isoform usage in 10/11 cohorts. We interrogated the effect of DNA methylation on APA isoform usage at individual cytosines across the majority of gene (chr7:774633-826027). We see a strong correlation spanning chr7:807596-809109 in BLCA and BRCA, as well as in the merged set (Figure 4A). All cohorts except for KIRC show a similar correlation (Figure S3A; Table S6). The cytosines preceding and following the identified region show no correlation. We summarized the data for the probes across the entire gene and found strong R2 and p-values at a restricted region of the gene (Figure 4B).

In order to examine the utility of DNA methylation and isoform usage as cancer biomarkers we analyzed available tumor stage data (Figure S3B&C). Several cancer cohorts showed a trend in *HEATR2* isoform usage across tumor stages. In particular, LUSC showed a significant difference between stage 4 and stage 1 (Figure S3B). The same appears true for DNA methylation. Several cancers show distinct trends in cytosine methylation across tumor stages. Often, this trend appears to be equivalent across the four analyzed cytosines (Figure S3C).

We next assessed the correlation between DNA methylation and isoform usage in all 546 candidate genes across the 11 cancer subtypes and the merged set (Figure 4C). 384 genes showed a correlation in at least one cancer subtype. Seven genes exhibited correlations across multiple cancer subtypes and of these *HEATR2* showed the best correlation (Figure S4D). Importantly, this remarkable correlation aligns with our differentially methylated CGI where we have observed differential binding of CTCF and RAD21 (Figure 4B). Additionally, PhastCon scores show strong sequence conservation, despite the fact that this region does not align with an exon.

**Discussion**

In the absence of DNA methylation, CTCF binds DNA and recruits the cohesin complex, stalling Pol2 and increasing proximal isoform production. DNA methylation prohibits CTCF binding, promoting distal isoform production (Figure 5). CTCF and the cohesin complex form DNA loops to establish topologically associating domains and define chromatin boundaries (Phillips and Corces, 2009; Song and Kim, 2017; Wutz *et al.*, 2017). In our mechanism a DNA loop may form that stalls Pol2 and promotes proximal isoform production in the demethylated context (Figure 5). RAD21 depletion eliminates chromatin loops after six hours of IAA treatment (Rao *et al.*, 2017). The loss of DNA loops may contribute to the observed distal isoform recovery. Further studies into the candidate genes’ chromatin structure will provide additional insight into structural aspects of our mechanism. Nonetheless, our current study demonstrates that DNA methylation, CTCF, and the cohesin complex cooperate to regulate transcriptome diversity.

Misregulated APA occurs in cancer with studies suggesting a trend towards global 3′UTR shortening (Di Giammartino, Nishida and Manley, 2011; Huang *et al.*, 2018). Shortened 3′UTRs drive tumorigenesis in many ways. Oncogenes, such as *IMP-1*, undergo APA-mediated shortening, thereby avoiding miRNA-mediated degradation (Mayr and Bartel, 2009). Conversely, shortened 3′UTRs of tumor suppressor competing endogenous RNAs increased miRNA-mediated degradation of tumor suppressors (Park *et al.*, 2018). Additionally, intronic APA drives tumorigenesis by generating truncated proteins (Lee *et al.*, 2018). Global patterns of APA vary between cancers; lengthened 3′UTRs are also seen in cancer (Fu *et al.*, 2011; Morris *et al.*, 2012). Previous research into APA regulation focused almost exclusively on concentrations of *trans*-acting factors (Takagaki and Manley, 1998; Tian and Manley, 2013; Nimura *et al.*, 2016; Huang *et al.*, 2018). This cannot account for all APA. Understanding the mechanisms that dysregulate APA will provide critical insights to its regulation in cancer. Genome-wide methylation is similarly dysregulated during carcinogenesis. For example, promoter hypomethylation activates oncogenes; protease urokinase is overexpressed and hypomethylated in breast and prostate cancers (Jaenisch and Bird, 2003; Ehrlich, 2009). Moreover, promoter hypermethylation silences tumor suppressors: Von Hippel-Lindau and CDKN2A are promoter hypermethylated in renal and colorectal cancers (Baylin and Jones, 2011; Shima *et al.*, 2011). Beyond the promoter, gene body methylation is a therapeutic target in cancer (Yang *et al.*, 2014).

DAC treatment, an established epigenetic cancer therapy, reactivates genes by promoter demethylation and represses genes by gene body demethylation (Momparler, 2005; Yang *et al.*, 2014). DAC treatment induces remission in several cancers (Malik and Cashen, 2014). However, the side effects of global demethylation are challenging to predict without an exhaustive understanding of DNA methylation’s regulatory effects. Our mechanism indicates that we should consider not only the transcriptional activity caused by demethylation but also the erasure of regulation. Methylation-regulated APA may identify therapeutic targets in cancer. NFY may have widespread effects on carcinogenesis as many of its targets are upregulated in cancer (Gurtner, Manni and Piaggio, 2017; Zambelli and Pavesi, 2017). The reasons behind these changes are currently unidentified. Methylation-regulated APA presents a new avenue of research as to NFYA’s role in carcinogenesis. HEATR2 has not previously been associated with cancer. However, there is a remarkably strong correlation between methylation of *HEATR2* and its isoform usage. As HEATR2 is important for ciliary motility, it may be involved in the motility of cancer cells (Diggle *et al.*, 2014; Horani *et al.*, 2018). These are examples of therapeutic insights that may arise from studying methylation-regulated APA.

Despite their prevalence in carcinogenesis, DNA methylation and APA are not pathogenic by nature. Rather, they are necessary for everyday regulation of gene expression and transcriptome diversity. Promoter DNA methylation is necessary for gene imprinting and X chromosome-inactivation and stabilizes gene expression by creating silent chromatin states (Jaenisch and Bird, 2003). Although gene bodies are ancient methylation targets, the effects of genic DNA methylation are unclear (Zemach and Zilberman, 2010). Methylation at centromeres may stabilize chromosomes and suppress transposable elements (Moarefi and Chédin, 2011). Intragenic methylation may silence alternative transcription start sites or affect the efficiency of Pol2 (Shenker and Flanagan, 2012). Approximately 22% of splicing appears to be regulated by exonic DNA methylation (Lev Maor, Yearim and Ast, 2015; Shayevitch *et al.*, 2018). APA is similarly prevalent in normal development. More than 50% of human genes produce multiple isoforms via APA (Tian *et al.*, 2005). For example, CD47 localizes to the cell surface with a long 3′UTR or the endoplasmic reticulum with a short 3′UTR (Berkovits and Mayr, 2015). Embryonic tissues express shorter 3′ UTRs than differentiated tissues, and between differentiated tissues, there is cell-specific expression of 3′ UTRs (Mayr, 2017). Additionally, intronic polyadenylation diversifies the immune cell transcriptomes (Singh *et al.*, 2018). These are vital aspects of development. Our mechanism unites DNA methylation, alternative cleavage and polyadenylation, and chromatin dynamics, all of which are misregulated in cancer. We show that DNA methylation promotes proximal pAS usage and that depletion of RAD21 recovers the distal APA isoform. This mechanism will contribute to further understanding of APA regulation and DNA methylation in development.

Baylin, S. B. and Jones, P. A. (2011) ‘A decade of exploring the cancer epigenome — biological and translational implications’, *Nature Reviews Cancer*. Nature Publishing Group, 11(10), pp. 726–734. doi: 10.1038/nrc3130.

Bell, A. C. and Felsenfeld, G. (2000) ‘Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene’, *Nature*. Nature Publishing Group, 405(6785), pp. 482–485. doi: 10.1038/35013100.

Berkovits, B. D. and Mayr, C. (2015) ‘Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization’, *Nature*. Nature Publishing Group, 522(7556), pp. 363–367. doi: 10.1038/nature14321.

Bhasin, J. M. *et al.* (2015) ‘Methylome-wide Sequencing Detects DNA Hypermethylation Distinguishing Indolent from Aggressive Prostate Cancer.’, *Cell reports*. NIH Public Access, 13(10), pp. 2135–46. doi: 10.1016/j.celrep.2015.10.078.

Bowman, E. A. and Kelly, W. G. (2014) ‘RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases’, *Nucleus*. Taylor & Francis, 5(3), p. 224. doi: 10.4161/NUCL.29347.

Ciark, S. J. *et al.* (1994) ‘High sensitivity mapping of methylated cytosines’, *Nucleic Acids Research*, 22(15), pp. 2990–2997. Available at: https://oup.silverchair-cdn.com/oup/backfile/Content\_public/Journal/nar/22/15/10.1093/nar/22.15.2990/2/22-15-2990.pdf?Expires=1502997906&Signature=G0YhODRM1pmJzFyOiACertH5O3DOHnyMBiyrR8G5o3J6st6gJf5KN5SdsPWr9kmqTyBQ3Hbd4~vCzl3oQ67VIaPwPiGn~e-HF5HpOfdXuo16ZK-h6zxtkqUkQhE9ECfT0-RwgW8NJKIeMllHijmKzJAOoVoeAsj8M2AVya33djRjfBhQsHQulDStWb0MwqurElqHNcFnUYy3GwdYzSNzmn7KPnFSHWYmeYydOU5oQudXUHyNLpKD6DizMCAsqQsI4IgHYJ94NGmowMGvnTtrlOxdUR7PY25nyhE4SuqCdZVIII-ltogR8sw76yfZX55A~yc35j1Mlb5GTOh1oFQivQ\_\_&Key-Pair-Id=APKAIUCZBIA4LVPAVW3Q (Accessed: 16 August 2017).

Creyghton, M. P. *et al.* (2010) ‘Histone H3K27ac separates active from poised enhancers and predicts developmental state.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(50), pp. 21931–6. doi: 10.1073/pnas.1016071107.

Diggle, C. P. *et al.* (2014) ‘HEATR2 plays a conserved role in assembly of the ciliary motile apparatus.’, *PLoS genetics*. Public Library of Science, 10(9), p. e1004577. doi: 10.1371/journal.pgen.1004577.

Ehrlich, M. (2009) ‘DNA hypomethylation in cancer cells.’, *Epigenomics*. NIH Public Access, 1(2), pp. 239–59. doi: 10.2217/epi.09.33.

Elkon, R., Ugalde, A. P. and Agami, R. (2013) ‘Alternative cleavage and polyadenylation: extent, regulation and function’, *Nature Reviews Genetics*. Nature Publishing Group, 14(7), pp. 496–506. doi: 10.1038/nrg3482.

Feng, S. *et al.* (2010) ‘Conservation and divergence of methylation patterning in plants and animals.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(19), pp. 8689–94. doi: 10.1073/pnas.1002720107.

Fu, Y. *et al.* (2011) ‘Differential genome-wide profiling of tandem 3′ UTRs among human breast cancer and normal cells by high-throughput sequencing’, *Genome Research*, 21(5), pp. 741–747. doi: 10.1101/gr.115295.110.

Ghosh, S. *et al.* (2010) ‘Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues.’, *Epigenetics*. Taylor & Francis, 5(6), pp. 527–38. doi: 10.4161/EPI.5.6.12228.

Di Giammartino, D. C., Nishida, K. and Manley, J. L. (2011) ‘Mechanisms and consequences of alternative polyadenylation.’, *Molecular cell*. NIH Public Access, 43(6), pp. 853–66. doi: 10.1016/j.molcel.2011.08.017.

Gruber, A. R. *et al.* (2014) ‘Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors.’, *Wiley interdisciplinary reviews. RNA*. Wiley-Blackwell, 5(2), pp. 183–96. doi: 10.1002/wrna.1206.

Gurtner, A., Manni, I. and Piaggio, G. (2017) ‘NF-Y in cancer: Impact on cell transformation of a gene essential for proliferation’, *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. Elsevier, 1860(5), pp. 604–616. doi: 10.1016/J.BBAGRM.2016.12.005.

He, X.-J., Chen, T. and Zhu, J.-K. (2011) ‘Regulation and function of DNA methylation in plants and animals.’, *Cell research*. Nature Publishing Group, 21(3), pp. 442–65. doi: 10.1038/cr.2011.23.

Horani, A. *et al.* (2012) ‘Whole-exome capture and sequencing identifies HEATR2 mutation as a cause of primary ciliary dyskinesia.’, *American journal of human genetics*. Elsevier, 91(4), pp. 685–93. doi: 10.1016/j.ajhg.2012.08.022.

Horani, A. *et al.* (2018) ‘Establishment of the early cilia preassembly protein complex during motile ciliogenesis.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, p. 201715915. doi: 10.1073/pnas.1715915115.

Huang, J. *et al.* (2018) ‘Suppression of cleavage factor Im 25 promotes the proliferation of lung cancer cells through alternative polyadenylation’, *Biochemical and Biophysical Research Communications*. Academic Press, 503(2), pp. 856–862. doi: 10.1016/J.BBRC.2018.06.087.

Jaenisch, R. and Bird, A. (2003a) ‘Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals’, *Nature Genetics*. Nature Publishing Group, 33(3s), pp. 245–254. doi: 10.1038/ng1089.

Jaenisch, R. and Bird, A. (2003b) ‘Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals’, *Nature Genetics*. Nature Publishing Group, 33(3s), pp. 245–254. doi: 10.1038/ng1089.

Jereb, S. *et al.* (2018) ‘Differential 3′ Processing of Specific Transcripts Expands Regulatory and Protein Diversity Across Neuronal Cell Types’, *bioRxiv*. Cold Spring Harbor Laboratory, p. 245886. doi: 10.1101/245886.

Lee, B.-S. *et al.* (2016) ‘Antibody Production with Synthetic Peptides’, in. Humana Press, New York, NY, pp. 25–47. doi: 10.1007/978-1-4939-6352-2\_2.

Lee, J. Y. *et al.* (2007) ‘PolyA\_DB 2: mRNA polyadenylation sites in vertebrate genes.’, *Nucleic acids research*. Oxford University Press, 35(Database issue), pp. D165-8. doi: 10.1093/nar/gkl870.

Lee, S.-H. *et al.* (2018) ‘Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia’, *Nature*. Nature Publishing Group, 561(7721), pp. 127–131. doi: 10.1038/s41586-018-0465-8.

Lev Maor, G., Yearim, A. and Ast, G. (2015) ‘The alternative role of DNA methylation in splicing regulation.’, *Trends in genetics : TIG*. Elsevier, 31(5), pp. 274–80. doi: 10.1016/j.tig.2015.03.002.

Li, E., Bestor, T. H. and Jaenisch, R. (1992) ‘Targeted mutation of the DNA methyltransferase gene results in embryonic lethality’, *Cell*, 69(6), pp. 915–926. doi: 10.1016/0092-8674(92)90611-F.

Lianoglou, S. *et al.* (2013) ‘Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression.’, *Genes & development*. Cold Spring Harbor Laboratory Press, 27(21), pp. 2380–96. doi: 10.1101/gad.229328.113.

Malik, P. and Cashen, A. F. (2014) ‘Decitabine in the treatment of acute myeloid leukemia in elderly patients.’, *Cancer management and research*. Dove Press, 6, pp. 53–61. doi: 10.2147/CMAR.S40600.

Masamha, C. P. *et al.* (2014) ‘CFIm25 links alternative polyadenylation to glioblastoma tumour suppression’, *Nature*. Nature Publishing Group, 510(7505), pp. 412–416. doi: 10.1038/nature13261.

Mayr, C. (2017) ‘Regulation by 3′-Untranslated Regions’, *Annual Review of Genetics*. Annual Reviews , 51(1), pp. 171–194. doi: 10.1146/annurev-genet-120116-024704.

Mayr, C. and Bartel, D. P. (2009) ‘Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.’, *Cell*. NIH Public Access, 138(4), pp. 673–84. doi: 10.1016/j.cell.2009.06.016.

Moarefi, A. H. and Chédin, F. (2011) ‘ICF Syndrome Mutations Cause a Broad Spectrum of Biochemical Defects in DNMT3B-Mediated De Novo DNA Methylation’, *Journal of Molecular Biology*. Academic Press, 409(5), pp. 758–772. doi: 10.1016/J.JMB.2011.04.050.

Momparler, R. L. (2005) ‘Pharmacology of 5-Aza-2′-deoxycytidine (decitabine)’, *Seminars in Hematology*. W.B. Saunders, 42, pp. S9–S16. doi: 10.1053/J.SEMINHEMATOL.2005.05.002.

Morris, A. R. *et al.* (2012) ‘Alternative Cleavage and Polyadenylation during Colorectal Cancer Development’, *Clinical Cancer Research*. American Association for Cancer Research, 18(19), pp. 5256–5266. doi: 10.1158/1078-0432.CCR-12-0543.

Natsume, T. *et al.* (2016) ‘Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors’, *Cell Reports*. Cell Press, 15(1), pp. 210–218. doi: 10.1016/J.CELREP.2016.03.001.

Nimura, K. *et al.* (2016) ‘Regulation of alternative polyadenylation by Nkx2-5 and Xrn2 during mouse heart development.’, *eLife*. eLife Sciences Publications, Ltd, 5. doi: 10.7554/eLife.16030.

Ogorodnikov, A., Kargapolova, Y. and Danckwardt, S. (2016) ‘Processing and transcriptome expansion at the mRNA 3′ end in health and disease: finding the right end.’, *Pflugers Archiv : European journal of physiology*. Springer, 468(6), pp. 993–1012. doi: 10.1007/s00424-016-1828-3.

Okano, M. *et al.* (1999) ‘DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development’, *Cell*, 99(3), pp. 247–257. doi: 10.1016/S0092-8674(00)81656-6.

Park, H. J. *et al.* (2018) ‘3′ UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk’, *Nature Genetics*. Nature Publishing Group, 50(6), pp. 783–789. doi: 10.1038/s41588-018-0118-8.

Phillips, J. E. and Corces, V. G. (2009) ‘CTCF: master weaver of the genome.’, *Cell*. NIH Public Access, 137(7), pp. 1194–211. doi: 10.1016/j.cell.2009.06.001.

Rädle, B. *et al.* (2013) ‘Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture.’, *Journal of visualized experiments : JoVE*. MyJoVE Corporation, (78). doi: 10.3791/50195.

Rao, S. S. P. *et al.* (2017) ‘Cohesin Loss Eliminates All Loop Domains.’, *Cell*. Elsevier, 171(2), pp. 305-320.e24. doi: 10.1016/j.cell.2017.09.026.

Renda, M. *et al.* (2007) ‘Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci.’, *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 282(46), pp. 33336–45. doi: 10.1074/jbc.M706213200.

Rhee, I. *et al.* (2000) ‘CpG methylation is maintained in human cancer cells lacking DNMT1’, *Nature*. Nature Publishing Group, 404(6781), pp. 1003–1007. doi: 10.1038/35010000.

Rhee, I. *et al.* (2002) ‘DNMT1 and DNMT3b cooperate to silence genes in human cancer cells’, *Nature*. Nature Publishing Group, 416(6880), pp. 552–556. doi: 10.1038/416552a.

Shayevitch, R. *et al.* (2018) ‘The importance of DNA methylation of exons on alternative splicing.’, *RNA (New York, N.Y.)*. Cold Spring Harbor Laboratory Press, 24(10), pp. 1351–1362. doi: 10.1261/rna.064865.117.

Shenker, N. and Flanagan, J. M. (2012a) ‘Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research.’, *British journal of cancer*. Nature Publishing Group, 106(2), pp. 248–53. doi: 10.1038/bjc.2011.550.

Shenker, N. and Flanagan, J. M. (2012b) ‘Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research.’, *British journal of cancer*. Nature Publishing Group, 106(2), pp. 248–53. doi: 10.1038/bjc.2011.550.

Shima, K. *et al.* (2011) ‘Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review.’, *International journal of cancer*. NIH Public Access, 128(5), pp. 1080–94. doi: 10.1002/ijc.25432.

Singh, I. *et al.* (no date) ‘Widespread intronic polyadenylation diversifies immune cell transcriptomes’. doi: 10.1038/s41467-018-04112-z.

Song, F. *et al.* (2005) ‘Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 102(9), pp. 3336–41. doi: 10.1073/pnas.0408436102.

Song, S.-H. and Kim, T.-Y. (2017) ‘CTCF, Cohesin, and Chromatin in Human Cancer.’, *Genomics & informatics*. Korea Genome Organization, 15(4), pp. 114–122. doi: 10.5808/GI.2017.15.4.114.

Suzuki, M. M. and Bird, A. (2008) ‘DNA methylation landscapes: provocative insights from epigenomics’, *Nature Reviews Genetics*. Nature Publishing Group, 9(6), pp. 465–476. doi: 10.1038/nrg2341.

Takagaki, Y. and Manley, J. L. (1998) ‘Levels of Polyadenylation Factor CstF-64 Control IgM Heavy Chain mRNA Accumulation and Other Events Associated with B Cell Differentiation’, *Molecular Cell*. Cell Press, 2(6), pp. 761–771. doi: 10.1016/S1097-2765(00)80291-9.

Tian, B. *et al.* (2005) ‘A large-scale analysis of mRNA polyadenylation of human and mouse genes.’, *Nucleic acids research*. Oxford University Press, 33(1), pp. 201–12. doi: 10.1093/nar/gki158.

Tian, B. and Manley, J. L. (2013) ‘Alternative cleavage and polyadenylation: the long and short of it.’, *Trends in biochemical sciences*. NIH Public Access, 38(6), pp. 312–20. doi: 10.1016/j.tibs.2013.03.005.

Weber, M. *et al.* (2007) ‘Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome’, *Nature Genetics*. Nature Publishing Group, 39(4), pp. 457–466. doi: 10.1038/ng1990.

Wutz, G. *et al.* (2017) ‘Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins.’, *The EMBO journal*. European Molecular Biology Organization, 36(24), pp. 3573–3599. doi: 10.15252/embj.201798004.

Xu, C. and Zhang, J. (2018) ‘Alternative Polyadenylation of Mammalian Transcripts Is Generally Deleterious, Not Adaptive’, *Cell Systems*. Cell Press. doi: 10.1016/J.CELS.2018.05.007.

Yang, X. *et al.* (2014) ‘Gene body methylation can alter gene expression and is a therapeutic target in cancer.’, *Cancer cell*. NIH Public Access, 26(4), pp. 577–90. doi: 10.1016/j.ccr.2014.07.028.

Yu, M. *et al.* (2006) ‘Alternative 3′ UTR polyadenylation of Bzw1 transcripts display differential translation efficiency and tissue-specific expression’, *Biochemical and Biophysical Research Communications*. Academic Press, 345(1), pp. 479–485. doi: 10.1016/J.BBRC.2006.04.113.

Zambelli, F. and Pavesi, G. (2017) ‘Genome wide features, distribution and correlations of NF-Y binding sites’, *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. Elsevier, 1860(5), pp. 581–589. doi: 10.1016/J.BBAGRM.2016.10.007.

Zemach, A. and Zilberman, D. (2010) ‘Evolution of Eukaryotic DNA Methylation and the Pursuit of Safer Sex’, *Current Biology*. Cell Press, 20(17), pp. R780–R785. doi: 10.1016/J.CUB.2010.07.007.

Zheng, D. *et al.* (2018) ‘Cellular stress alters 3′UTR landscape through alternative polyadenylation and isoform-specific degradation’, *Nature Communications*. Nature Publishing Group, 9(1), p. 2268. doi: 10.1038/s41467-018-04730-7.