# Discovery of N6-methyladenosine on R-loops and elucidating its role in their regulation

**Abdulkadir H Abakir, MSci Hons, MSc**

**Division of Cancer and Stem Cells**
**Faculty of Medicine**
**University of Nottingham**

A thesis submitted for the degree of Doctor of Philosophy
At the University of Nottingham
July 2020
Word count: 50390

**Abstract**

Amongst the hundreds of known chemical modification of RNA, N6-methyladenosine (m6A) is the most abundant internal modification found on messenger RNAs. Despite this, the biological function of m6A remained largely enigmatic, in part due to earlier studies suggesting its static nature. The discovery of enzymes capable of reversing m6A demonstrated its dynamic nature. This was complemented by development of methodologies that allowed transcriptome-wide analysis of m6A distribution in eukaryotic organisms from *yeast* to human, with the subsequent implication of m6A in almost all steps of RNA metabolism. Meanwhile, N6-methyldeoxyadenine (6mA), its DNA counterpart that, for decades, was thought to be restricted to prokaryotic genomes and single cell eukaryotes, has recently been detected in multi-cellular eukaryotes *C. elegans, Drosophila* and green algae. However, due to contradictory reports regarding its detection, the presence and potential biological function of 6mA in mammalian genomes was unclear. In search of 6mA in human pluripotent stem cell genome, this thesis corroborated other studies in failing to detect 6mA in human cells. Nonetheless, this thesis describes the discovery of m6A presence on the RNA component of R-loops, the three-stranded nucleic acid structures consisting of RNA:DNA hybrid and a displaced single stranded DNA, that can confer vulnerability to genotoxic attacks. Through analysis of cell-cycle resolved R-loops, this thesis reveals the dynamic nature of m6A modified RNA:DNA hybrids, notably, with their higher accumulation in S and G2/M cell cycle phases compared to G0/G1 cells. Moreover, this thesis uncovers direct interaction of the m6A reader, YTHDF2 with R-loops, both *in vitro* and *in vivo*. Finally, this thesis uncovers a role for both m6A and YTHDF2 in regulating the stability of R-loops, as depletion of either the m6A writer, METTL3 or YTHDF2 resulted in genome-wide R-loops accumulation. Moreover, persistent loss of YTHDF2 led to cell growth retardation, increased accumulation of RNA:DNA hybrids and extensive DNA damage in R-loop-containing loci. Taken together, the findings of this thesis uncover a role for m6A and YTHDF2 in safeguarding the genome by limiting toxic accumulation of R-loops.

**Acknowledgments**

I wish to express my deepest gratitude to my supervisor, Dr Alexey Ruzov for the opportunity to join his lab for my PhD. I am truly indebted for his supervision, mentorship, collegiality, and unwavering support, in the ups and downs, both inside and outside the lab. It has been a real privilege to complete my PhD under his exemplary supervision, and scientific leadership, but also for his humanity and decency. I would also like to thank all my of lab colleagues, both past and present, including Lara, Ashley, Maria, Paige, and in particular, those I had the privilege to bench supervise their MSc thesis, Fahad and Liam.

I would also like to extend my sincerest thanks to Professor David Bates, who offered to be my second supervisor, for his constant encouragements, and protection from some of the none-sense around. Although we did not always agree, I really appreciate his openness. I would like to extend my thanks to members of the division, especially Dr Alexander Kondrashov for all the insightful scientific discussions.

I would like to express my profound gratitude to Professor Arne Klungland for hosting my research visits, his scientific guidance and exemplary support without which my project would not have been complete. I would also like to extend my thanks to all members of the Klungland group, especially Miaomiao Li and Sonia Perez for their enthusiastic help in all the experiments we performed in Oslo. I am immensely thankful to Professor Magnar Bjørås for offering me a stunning accommodation during my second visit to Oslo, his scientific encouragements and all the insightful scientific discussions. I would like to extend specialthanks to Professor Bjørn Dalhus for all the help in demonstrating the m6A dependent YTHDF2 interaction with RNA:DNA hybrids. Professor Klungland, his team and collaborators that together made my stay in Oslo not only fun but extremely fruitful.

My PhD project would not have been possible without the wonderful help and guidance of Professor Daniel Gackowski, whose amazing team not only key to the quantification of various DNA and RNA modifications but crucially, they helped us demonstrate the presence and quantification of m6A on R-loops. I am deeply honoured to have an ongoing collaboration with his lab.

I would like to thank Dr Natalia Gromak at the University of Oxford for all the help in independently validating the presence of m6A on R-loops as well as establishing the interaction of key m6A reader proteins with R-loops. I am especially grateful for the opportunity to join her lab for my MRC placement, unwavering support and the insightful

discussions that truly deepened my understanding of transcription and R-loop biology. I am immensely thankful to all the members of her lab (Agnese, Chiara, and Flavia) as well as all the members of Professors Norbury, Murphy, Proudfoot labs and the rest of the Dunn school, for making my stay in Oxford thoroughly enjoyable.

I would like to thank Dr Martin Gering and members of his team (Peter, Jasmin, Jo, Fatima and Gawahir) for welcoming me into their lab during the revision of our manuscript, and for allowing me unfettered access to all their reagents and lab equipments for my experiments. I would also like to thank Dr Gering for his scientific guidance, honest advice and amazing humour (from science to football and back) that kept me coming back for more and without which my PhD experience would have been terrible.

I would like to also thank Professor Andrew Jackson for providing me access to his tissue-culture facility and lab resources, without which I would have struggled. I would also like to extend my thanks to members of his lab (Darren, Jodie, and Luke) for willing to share their lab equipments.

Thank you to Dr William Brown for offering me access to his DT40 cell lines and reagents for culturing them during the early years of my PhD. I am particularly grateful for his open-door policy with me – which led to fascinating discussions about potential directions of my PhD project. I am truly grateful for his advice and support.

None of the genome-wide mapping experiments would have been possible without the heroic efforts and support of two amazing teams, Deep-seq and ADAC-Nottingham. I would therefore like to thank both past and present members of deep-seq (Sunir, Fei, Nadine, Vicki, Chris, Matt Carlile and Matt Loose) and ADAC (Tom Giles and Richard Emes). Their dedication, professionalism and efficiency contributed to publishing our results in Nature Genetics.

I would like to thank members of the doctoral programs committee (DPC) for electing me along with Joe Howie to represent school of medicine students across the 13 divisions on the DPC. It has been a real privilege to play a role in heling to improve the experience of so many students in the Medical School. I would like to particularly thank Dr Karen Robinson and Dr Lorraine Pennington for encouraging me to take-up this role.

I would also like to thank the MRC IMPACT DTP board both past (especially the legendary Steve West and David Lambert) and present for all the funding support throughout my PhD,

without which none of the experiments both here and abroad would have been possible. I would also like to thank the MRC IMPACT DTP students for electing me to represent them on the IMAPCT management board. I am proud of the many successes we have made.

I would also like to thank my former tutor, Dr Paul Scotting for encouraging me to pursue a PhD, his unconditional support throughout the years. I owe so much of gratitude to his kindness.

Special thanks to these SLIM facility members, Robert Markus, Seema Rajani, Denise McLean, and Tim Self for their training, ad-hoc support, and collaborations. I am also deeply grateful for their outstanding support in making the outreach activity we organised for a local primary school a success.

I would like to pay special thanks to my thesis defence committee, Professor Reinhard Stoger from the University of Nottingham, and Professor Eric Greer from Harvard University for examining and most importantly approving of my PhD defence. In addition, I would like to thank Professor Martin Gering coordinating and chair the thesis viva.

Most importantly, I would like to thank every member of my family, especially my parents, wife, and wonderful kids – I am, because of you.

**Table of contents**

Abstract.....2
Acknowledgements.....3
Table of Contents.....6
List of Figures.....12
List of Tables.....13
Abbreviations .....15

**Chapter 1:**

**Introduction**.....18

**1.1. Methylation of deoxyadenosine and its biological roles**
1.1.1. Various nucleic acids are methylation substrates.....19
1.1.2. Deoxyadenosine methylation in prokaryotes.....20
1.1.3. Detection of deoxyadenosine methylation in eukaryotic systems.....20
1.1.4. Writers and erasers of deoxyadenosine methylation .....21
1.1.5. Biological consequences of deoxyadenosine methylation in eukaryotes .....22
1.1.6. Contradictory reports of deoxyadenosine methylation in mammals.....23
**1.2. N6-methyladenosine (m6A) in RNA**
1.2.1. Methylation of RNA adenosine is dynamic.....23
1.2.2. Writers of m6A .....25
1.2.3. Erasers of m6A.....27
1.2.4. Readers of m6A.....28
**1.3.m6A deposition impacts multiple steps of RNA biogenesis**
1.3.1. Role of m6A in RNA Stability and degradation.....30
1.3.2. The role of m6A in regulating RNA splicing.....31
1.3.3. Role of m6A in nuclear RNA export.....31
1.3.4. The role of m6A in translation.....32
**1.4.Cellular and physiological functions of m6A modified RNA**
1.4.1. The role of m6A in meiosis .....32
1.4.2. The role of m6A during stress response .....33
1.4.3. The Role of m6A in cellular differentiation and development.....34
1.4.4. The role of m6A in disease.....36
**1.5.RNA:DNA hybrids (R-loops)**
1.5.1. Definitions and potential origins.....38
1.5.2. Consequences of RNA:DNA hybrid formation.....40
1.5.3. The role of RNA:DNA hybrids in DNA replication.....40
1.5.4. The role of RNA:DNA hybrids in Ig class switch recombination.....41
1.5.5. The role of RNA:DNA hybrids in transcriptional regulation.....42
1.5.6. RNA:DNA hybrids and genome instability.....44
1.5.7. Mechanisms of RNA:DNA hybrid resolution.....46
1.5.8. Detection and mapping of RNA:DNA hybrids.....47

**1.6.Aims of the thesis**

1.6.1. Examine potential presence of N6-methyladeoxyadenosine (6mA) in the human genome…….49
1.6.2. Investigate the origin of chromatin-associated 6mA/m6A immunostaining signal in human cells……49
1.6.3. Investigate the genome-wide distribution of m6A marked RNA:DNA hybrids in human cells……...50
1.6.4. Interrogate the relationship between m6A marked RNA:DNA hybrids and transcription……..……50
1.6.5. Investigate the relationship between histone modifications and m6A marked RNA:DNA hybrids in human pluripotent stem cells……50
1.6.6. Investigate the cell-cycle dynamics of m6A marked RNA:DNA hybrids in human pluripotent stem cells……51
1.6.7. Investigate the biological function of m6A on RNA:DNA hybrids……...51
1.7.List of publications……...52

**Chapter 2: Materials and Methods**

2.1.**Materials**:
2.1.1. Cell lines and mouse tissues……55
2.1.2. Antibodies……55
2.1.3. Immunofluorescence reagents ……..……56
2.1.4. Cell culture reagents……57
2.1.5. Flow cytometry materials and reagents to sort human cells into G0/G1, S and G2/M phases of the cell cycle……59
2.1.6. Chromatin immunoprecipitation reagents……59
2.1.7. Materials and reagents for m6A and S9.6 DRIP as well as m6A RIP……60
2.1.8. Oligonucleotides……...62
2.1.9. siRNAs, transfection reagents, bacteria, plasmids……66
2.1.10. PCR, RT-PCR, and RT-qPCR……..66
2.1.11. Microscopes ……67
2.2.**Cell culture experiments**
2.2.1. Human pluripotent stem cells……...67
2.2.2. KaryoMAX Colcemid treatment……..68
2.2.3. Fluorescence-activated cell sorting (FACS)……..…68
2.2.4. Cell culture for cancer cell lines……69
2.2.4.1. LN18, LN229, and U87MG glioblastoma cell lines……69
2.2.4.2. HeLa cell culture……..69
2.2.4.3. HAP1 cell culture……69
2.2.4.4. Wild type and *ythdf2* knockout mouse neural stem cells……69
2.2.4.5. Growth curve of Wild type and YTHDF2 KO HAP1 cells……70
2.2.4.6. Short interference RNA knockdown of target genes……70
2.2.4.7. pEGFP-RNASEH1 plasmid transfections in HAP1 cells……70

2.3.**Molecular Biology Experiments**
2.3.1. Growth of pEGFP-RNASEH1 plasmid containing bacteria……71

2.3.2. Isolation of pEGFP-RNASEH1 plasmid……..71
2.3.3. Verification of pEGFP-RNASEH1 plasmid……71
2.3.4. Verification of Wild type and YTHDF2 KO cells.……71
2.3.5. Functional validation of YTHDF2 KO in HAP1 cells by immunostaining…..72

2.4.**Immunostaining experiments**
2.4.1. m6A immunocytochemistry in non-denaturing conditions…….……73
2.4.2. Co-immunostaining of m6A and 5-mC in denaturing conditions……73
2.4.3. Co-immunostaining of m6A and PH-3, 6mA and 5mC in no RNase condition, RNase A only and RNases A and H…….…….....74
2.4.4. Co-immunostaining of YTHDF2 and RNA:DNA hybrids (S9.6)…….....75
2.4.5. Co-immunostaining of YTHDF1 and PH-3 in hPSCs cells…….….76
2.4.6. Co-immunostaining of YTHDF2 and PH-3 in hPSCs…….….77
2.4.7. Co-immunostaining of HNRNPA2B1 and m6A in hPSCs……..…..77
2.4.8. Co-immunostaining of YTHDF2 and Ph-3 in WT and YTHDF2 KO cells….77
2.4.9. Immuno-competition assay for validation of anti-m6A antibody specificity...77
2.4.10. Immunostaining of 1-methyladenosine (m1A) in hPSCs……..78
2.4.11. Immunostaining of METTL3 KD hPSCs for phosphorylated γH2AX ……78
2.4.12. Immunostaining of HNRNPA2B1 KD hPSCs for phosphorylated γH2AX….78
2.4.13. Immunostaining of phosphorylated γH2AX in WT and YTHDF2 KO HAP1 Cells……..……78
2.4.14. Immunostaining of wild type and *ythdf2* KO mouse embryonic brain sections for phosphorylated γH2AX…….....79
2.5.**Microscopy and image analysis**
2.5.1. Widefield microscopy…….……..80
2.5.2. Confocal microscopy images……80
2.5.3. Imaging analysis……...……80
2.6.**Liquid Chromatography–Tandem Mass Spectrometry (LC–MS/MS)**
2.6.1. Preparation of cells for DNA and RNA isolation……..…80
2.6.2. Isolation of genomic DNA and RNA……81
2.6.3. Detection and quantification of m6A in genomic DNA and total RNA using LC-MS/MS …….……81
2.7.**Genome-wide mapping of m6A modified RNA:DNA hybrids in mammalian cells**
2.7.1. DNA isolation…….……..82
2.7.2. Quantification of genomic DNA……..……..…82
2.7.3. Sonication of genomic DNA……..…82
2.7.4. RNase H treatment of control samples for m6A and S9.6 DRIP……..…83
2.7.5. DNA adaptor ligation and sequencing library preparation for Illumina……83
2.7.6. S9.6 immunoprecipitation ……83
2.7.7. m6A RNA:DNA immunoprecipitation (m6A DRIP)……...…84
2.7.8. Library preparation and high throughput m6A and S9.6 DRIP sequencing …85
2.8.**Whole-transcriptome sequencing**
2.8.1. Total RNA extraction..……..87
2.8.2. RNA library preparation and sequencing..……87
2.9.**Bioinformatics analysis**
2.9.1. Analysis of m6A DIP- and S9.6DRIP-seq data…….……87

2.9.2. Bioinformatics: Comparison of the m6A DIP and S9.6 DRIP peaks datasets with genomic distribution of histone modifications in hPSCs……..……88
2.9.3. Analysis of gene expression for the m6A and S9.6 peaks-containing genes…89
2.9.4. The distribution of m6A marked RNA:DNA hybrids among different repetitive sequences…….…….90

2.10. **Validation of m6A DIP and S9.6 DRIP protocols**
2.10.1. Two-round RNA:DNA hybrid immunoprecipitations (S9.6 DRIP followed by m6A DRIP)……...90
2.10.2. Validation of m6A DRIP procedure: m6A RIP qRT-PCR following S9.6 DRIP…….93
2.10.3. Validation of m6A DIP and S9.6 DRIP using synthetic oligoes……93
2.10.4. Validation of m6A presence on RNA:DNA hybrids using stable isotope dilution ultra-performance liquid chromatography with tandem mass spectrometry (SID-UPLC-MS/MS) …….……..94
2.11. S9.6 DRIP slot-blot……...97
2.12. Chromatin Immunoprecipitation (ChIP)……...97
2.13. Electrophoretic Mobility-Shift Assay (EMSA)……98
2.14. Microscale thermophoresis……...……99
2.15. Statistics ……100

**Chapter 3: Results**

**3. N6-methyladenosine marks the RNA component of RNA:DNA hybrids (R-loops) in mammalian cells** ……102

3.1.m6A is detectable by immunochemistry and LC-MS/MS in hPSCs RNA……..103
3.2.Detection of an RNase A resistant m6A immunostaining signal in mitotic chromatin of various human cell lines……..…105
3.3.Validating the specificity of anti-m6A antibody with synthetic 6mdATP in hPSCs...107
3.4.Discovery of m6A presence on the RNA component of RNADNA hybrids (R-loops)……..108
3.5.Summary and conclusions…….….111

**Chapter 4: Results**

**4. Genome-wide analysis reveals majority of RNA:DNA hybrids in hPSCs are marked by m6A.**
4.1.Development of a new genome-wide technique for mapping the m6A-modified RNA:DNA hybrids….....113
4.2.Validation of the m6A DIP approach……..…113
4.3.Genomic distribution of m6A modified RNA:DNA hybrids……..…116
4.4.Analysis of m6A and S9.6 peak distribution across various genomic features……118
4.5.Summary and conclusions……..120

**Chapter 5: Results**

**5. Deposition of m6A on RNA:DNA hybrids correlate with transcription**
5.1.m6A and S9.6 peaks predominantly occur in transcribed genomic regions……....122

5.2.m6A and S9.6 DRIP-seq peaks strongly correlate with transcriptional activity of retroelements...……...123
5.3.m6A DIP and S9.6 DRIP peaks datasets do not overlap with majority of histone modifications in hPSCs……..….125
5.4.Summary and conclusions……..126

**Chapter 6: Results**

**6. m6A regulates cell-cycle dependent dynamics of RNA:DNA hybrids in hPSCs**
6.1. Cell-cycle dependent m6A deposition on RNA:DNA hybrids in human cells…..…129
6.2. Depletion of nuclear RNase H1 results in accumulation of m6A marked RNA:DNA hybrids in hPSCs……...…..133
6.3.Sequential S9.6 and m6A DRIP confirm enrichment of R-loops across multiple loci.135
6.4.Summary and conclusions…….……136

**Chapter 7: Results**

**7. m6A regulates the stability of RNA:DNA hybrids in human cells**
7.1.METTL3 depleted hPSCs exhibit cell-cycle depended accumulation of RNA:DNA hybrids……...…….138
7.2.METTL3 depleted hPSCs display genome-wide increase in m6A marked R-loops levels……..…139
7.3.m6A writers and readers are amongst proteins physically interacting with R-loops from human cell extracts ……..…..141
7.4.Summary and conclusion……142

**8. Dissecting the functional significance of YTHDF2 interaction with m6A-modified RNA:DNA hybrids.**
8.1.Cytoplasmic YTHDF2 localizes to chromatin of mitotic hPSCs……144
8.2.YTHDF2 co-localises with R-loops in *in vivo*……145
8.3.YTHDF2 interaction with synthetic RNA:DNA hybrids is m6A dependent……..147
8.4.YTHDF2 exhibit higher binding affinity towards m6A containing RNA:DNA hybrids compared to unmodified RNA:DNA hybrids……..…...148
8.5.Chromatin immunoprecipitation (ChIP) reveal interaction of YTHDF2 with R-loop-containing genomic loci……..149
8.6.Enrichment of YTHDF2 at R-loop-containing genomic loci in METTL3 and RNase H1 depleted hPSCs……...…150
8.7.YTHDF2 depletion leads to accumulation of RNA:DNA hybrids in hPSCs…….151
8.8.Summary and conclusions……..155

**9. Increased DNA damage and cell growth retardation: Biological consequences of YTHDF2 depletion.**

9.1.YTHDF2 and METTL3 depletion leads to elevated levels of H2AX phosphorylation in human and mouse cells……...157
9.2.YTHDF-2 depletion lead to cell growth retardation……...159
9.3.RNase H1 overexpression partially rescues increased γH2AX immunostaining signal of YTHDF2 KO cells……..…159
9.4.Summary and conclusions……..160
9.5.

**10. General discussion and future directions.**

10.1. Discovery of m6A on RNA:DNA hybrids…….….163
10.2. Genomic distribution of distribution of m6A marked RNA:DNA hybrids…165
10.3. Lack of correlation between m6A marked RNA:DNA hybrid and major histone modifications……...……...167
10.4. The cell-cycle as potential regulator of R-loop dynamics……..168
10.5. Modulation of m6A levels uncover its role in regulating R-loop stability…..169
**11. References** ……...173

**List of figures**

**Chapter 1.**

1.2.1. The dynamic and reversible nature of m6A RNA methylation

1.2.4. Domain architecture of human YTH domain-containing proteins

**Chapter 2**

2.4. Validation of YTHDF2 KO in HAP1 cells.

2.7.3. Visualization of a typical sonicated DNA fragments using the Agilent 2200 TapeStation system using D1000 reagents.

2.7.8. Workflow for m6A DRIP and S9.6 DRIP sequencing techniques

2.10.4. SID-UPLC-MS/MS analysis of hPSCs-derived nucleic acids released and retained upon RNase H treatment

**Chapter**

3.1. Detection of RNA m6A in hPSCs using immunochemistry and LC-MS/MS.

3.2. Mitotic chromatin of multiple human cell types exhibit RNase A resistant m6A immunostaining signal.

3.3. m6A immunostaining in hPSCs is specifically competed by 6mdATP but not by unmodified Deoxythymidine triphosphate.

3.4. m6A marks the RNA component of RNA:DNA hybrids in hPSCs.

**Chapter 4**

4.2. Validation of m6A DRIP using synthetic RNA and DNA substrates

4.3.1. Replicate correlation analysis for m6A and S9.6 DRIP-seq peaks

4.3.2. m6A is present on the majority of RNA:DNA hybrids in hPSCs.

4.4. Analysis of m6A and S9.6 peak distribution across distinct genomic features.

**Chapter 5**

5.1. The gene expression analysis of m6A and S9.6 peaks-containing genes

5.2. The distribution of m6A marked RNA:DNA hybrids among different repetitive sequences.

5.3. Comparison of the m6A DIP and S9.6 DRIP peaks datasets with genomic distribution of histone modifications in hPSCs

**Chapter 6**

6.1.1. FACS sorting of hPSCs and selection of RNA:DNA hybrid containing loci for analysis of their cycle-cycle.

6.1.2. Cell-cycle dependent dynamics of RNA:DNA hybrids

6.2. m6A DIP signal is increased upon RNase H1 knockdown in hPSCs

6.3. m6A is detectable on the RNA components of R-loops

**Chapter 7**

7.1. METTL3 depletion leads to accumulation of RNA:DNA hybrids in hPSCs.

7.2. METTL3 depletion leads to an increase in the numbers of m6A DIP peaks in hPSCs

7.3. m6 A writer and reader proteins interact with RNA:DNA hybrids.

**Chapter 8**

8.1. Immunostaining of hPSCs for m6A readers proteins.

8.2. YTHDF2 co-localizes with R-loops *in vivo*.

8.3. YTHDF2 interacts with synthetic m6A-marked RNA:DNA hybrids *in vitro*.

8.4. YTHDF2 exhibit higher binding affinity towards m6A containing RNA:DNA hybrids

8.5. ChIP-qPCR reveals YTHDF2 interacts with R-loop-containing loci *in vivo*.

8.6. Increased YTHDF2 recruitment to R-loop-containing genomic in RNase H1 or METTL3 depleted hPSCs.

8.7.1. YTHDF2 depletion leads to accumulation of R-loops, increased accretion of m6A on RNA:DNA hybrids

8.7.2. YTHDF2 depletion leads to accumulation of R-loops in human and mouse cells.

**Chapter 9**

9.1. YTHDF2 depletion leads to elevated levels of H2AX phosphorylation in human and mouse cells.

9.2. YTHDF2 depletion leads to cell growth retardation in mammalian cells

9.3. RNase H1 over-expression partially rescues the increased H2AX phosphorylation in YTHDF2 depleted HAP1 cells.

**List of tables:**

2.1.2. List of primary and secondary antibodies used for immunofluorescence (IF), Western blot (WB), ChIP and DRIP procedures.

2.1.3. Immunofluorescence reagents for mammalian cells, and paraffin embedded tissues.

2.1.4.1. Human pluripotent stem cell culture reagents. This includes reagents for both REBL-PAT hiPSCs and HUES7 hESCs.

2.1.4.2. Cell culture reagents for HeLa, LN-18 and U87MG cancer cell lines.

2.1.4.3. Cell culture reagents for the Wild type and YTHDF2 KO HAP1 cells.

2.1.5. Flow cytometry materials and reagents to sort human cells into G0/G1, S and G2/M phases of the cell cycle.

2.1.6. Chromatin immunoprecipitation reagents

2.1.7.1. Materials and reagents for m6A and S9.6 DRIP as well as m6A RIP

2.1.7.2. Buffers for m6A and S9.6 DRIP:

2.1.8.1. Primers used for RT-qPCR (For assessing the expression levels of the corresponding loci).

2.1.8.2. Primers used for verifying the genotype of Wild type and CRISPR cas9 mediated YTHDF2 KO HAP1 cells.

2.1.8.3. Primers used for investigating enrichment in m6A DRIP, S9.6 DRIP and ChIP experiments.

2.1.8.4. Oligonucleotides used for spike-in during m6A and S9.6 DRIP experiments

2.1.8.4. Oligonucleotides used in electrophoretic mobility-shift assays (EMSA).

2.1.8.5. Buffers used in EMSA

2.1.9. siRNAs, transfection reagents, bacteria, plasmids

2.1.10. PCR, RT-PCR, and RT-qPCR

2.1.11. Microscopes

**Abbreviations:**

DNA (deoxyribonucleic acid)

4mC (N4-cytosine)

5mC (5-methylcytosine)

4mC (N4-methylcytosine)

6mA (N6-methyldeoxyadenosine)

m6A (N6-methyladenosine)

m2G (N2-methylguanine)

m7G (N7-methylguanine)

N7MeG (N7-methyladenine)

O6mG (O6-methylguanine)

N3mT (N3-methylthymine)

O4mT (O4-methylthymine)

IS10 (bacterial transposable element)

IS50 (bacterial transposable element)

SAMe (S-adenosyl-L-methionine)

IME4 (inducer of meiosis 4)

MTA1c (ciliate 6mA methyltransferase)

METTL3 (methyltransferase 3)

METTL4 (methyltransferase 4)

METTL14 (methyltransferase 14)

METTL16 (methyltransferase 16)

WTAP (WT1 Associated Protein)

RBM15B (RNA Binding Motif Protein 15B)

RBM15 (RNA Binding Motif Protein 15)

ALKBH1 (AlkB Homolog 1)

ALKBH4 (AlkB Homolog 4)

ALKBH5 (AlkB Homolog 5)

H3K4me2 (H3 lysine 4 dimethyl)

N6MT1 (putative mammalian N-6 adenine-specific DNA methyltransferase 1)

DAMD (*Drosophila* 6mA methyltransferase)

TET (ten-eleven translocation)

NMAD-1 (*C. elegans* 6mA demethylase)

mRNA (messenger RNA)

3'UTR (3' untranslated region)

5'UTR (5' untranslated region)

TSS (Transcription start site)

RBM15 (RNA Binding Motif Protein 15)

FTO (Fat mass and obesity-associated protein)

5hmC (5-hydroxymethylcytosine)

5fC (5-formylcytosine)

5caC (5-carboxycytosine)

m6Am (N6,2-O-dimethyladenosine)

YTHDF1 (YTH domain-containing family protein 1)

YTHDF2 (YTH domain-containing family protein 2)

YTHDF3 (YTH domain-containing family protein 3)

YTHDC1 (YTH Domain Containing 1)

YTHDC2 (YTH Domain Containing 2)

HNRNPA2B1 (Heterogeneous Nuclear Ribonucleoprotein A2/B1)

HNRNPC (Heterogeneous Nuclear Ribonucleoprotein C)

HNRNPG (Heterogeneous Nuclear Ribonucleoprotein G)

PRRC2A (proline-rich and coiled-coil-containing protein 2A)

HRSP12 (heat-responsive protein 12)

eIF3 (eukaryotic translation initiation factor 3)

CSR (class switch recombination)

BER (base excision repair)

AID (activation-induced cytidine deaminase)

UNG (uracil DNA glycosylase)

APE1 (apurinic/apyrimidinic endonuclease 1)

RNase H (Ribonuclease H)

RNase A (Ribonuclease A)

Pol II (RNA polymerase II)

Pol III (RNA polymerase III)

THO (THO Complex)

HPR1 (Hyperrecombination protein 1)

MFT1 (a subunit of THO complex)

THP2 (a subunit of THO complex)

hPSCs (human pluripotent stem cells)

hiPSCs (human induced pluripotent stem cells)

hESCs (human embryonic stem cells)

WT (wild type)

KO (knockout)

# Chapter 1

# Introduction

# Chapter 1

## 1. Introduction

### 1.1. Methylation of deoxyadenosine and its biological roles

#### 1.1.1. Various nucleic acid bases are methylation substrates

Soon after the elegant demonstration that DNA serves as the hereditary genetic material (Avery et al., 1944), the existence of a modified nucleic acid base, 5-methylcytosine (5mC), was uncovered in a preparation of calf thymus (Hotchkiss, 1948), thus, preceding the elucidation of DNA double-helix structure (Watson and Crick, 1953). Despite its early identification, the appreciation of 5mC importance in generating phenotypic variation exclusive of changes in the DNA sequence, and, therefore its epigenetic (Waddington, 1959) significance, came much later (Holliday and Pugh, 1975). Through its predominant association with repressed chromatin, 5mC plays essential roles in a wide range of biological processes, including imprinting (Reik et al., 1987), X-chromosome inactivation (Lyon, 1961; Mohandas et al., 1981), transposon silencing (Walsh et al., 1998), embryonic development (Okano et al., 1999) and ageing (Wilson and Jones, 1983). Though 5mC remains the most studied DNA modification, the discovery of targeted methylation of the exocyclic nitrogen atom in position 6 of adenine bases (Dunn and Smith, 1955; Ratel et al., 2006) along with other modified nucleic acid bases (Sood et al., 2019) reflected the expanding lexicon of DNA modifications (Qifan Zhu et al., 2018). Notable modifications include, N4-methylcytosine (4mC) and N6-methyldeoxyadenosine (6mA) that comprise the restriction-modification system in prokaryote (Arber and Dussoix, 1962; Yu et al., 2015), whereby the host selectively modifies its genome in a defensive mode that allow targeting of unmodified foreign genome by restriction endonucleases (Ehrlich et al., 1987). With improvement in detection technologies, more than 40 types of chemically modified bases have been reported in DNA to date (Sood et al., 2019). Despite the diversity of DNA modifications, not all of

them are generated enzymatically, some may indicate DNA lesions such as N2-methylguanine, N7-methylguanine, N7-methyladenine, O6-methylguanine (6mG), N3-methylthymine, and O4-methylthymine (Sedgwick et al., 2007; Grosjean, 2009; Ito and Kuraoka, 2015; Kafer et al., 2016). As the mechanism by which some DNA modifications are generated remain enigmatic (Olinski et al., 2016), the very existence of some reported DNA modifications remain an open question (Schiffers et al., 2017; Stöger, 2017; O'Brown et al., 2019), never mind what function they might have (Chen et al., 2016).

### 1.1.2. Deoxyadenosine methylation in prokaryotes

First discovered in protists, the DAM (DNA adenine methylase) methyltransferase-generated, N6-methyl**deoxy**adenosine (6mA) is a pervasive DNA modification in prokaryotic genomes whose primary function is thought to distinguish self from invading DNA as part of the restriction-modification system (Dunn and Smith, 1955; Iyer et al., 2011; Vanyushin et al., 1970). In addition to its well documented role in host-genome defence, 6mA has been shown to regulate important biological processes, including bacterial DNA replication by mediating the DnaA binding at the oriC region (Boye et al., 2000; Reisenauer et al., 1999), replication associated mismatch repair (Hsieh, 2001), regulation of gene expression (Camacho et al., 2005; Low et al., 2001), and repression of IS10 and IS50 bacterial transposons by physically hindering RNA polymerase recruitment to their promoters (Roberts et al., 1985; Dodson and Berg, 1989). Collectively, these studies not only illustrate the essential functions of 6mA in the prokaryotic genome, but also motivated much of the recent excitement about the existence and potential function of 6mA in eukaryotic systems.

### 1.1.3. Detection of methylated adenine in eukaryotic systems.

Unlike prokaryotic DNA, eukaryotic genomes were initially thought to be devoid of 6mA (Vanyushin et al., 1970). The first compelling evidence of 6mA presence in eukaryotes was its discovery in transcriptionally active regions of *Tetrahymena* genome (Gorovsky et al.,

1973). This was subsequently followed by reports of 6mA in other ciliates, including *Paramecium aurelia* (Cummings et al., 1974), *Oxytricha fallax* (Rae and Spear, 1978) as well as *Stylonichia mytilius* (Ammermann et al., 1981; Gutiérrez et al., 2000). Comparative genomic analysis of DNA methyltransferases postulated the lateral transfer of adenine methylases to a variety of other eukaryotic systems, including the unicellular *Chlamydomonas reinhardtii* (Hattman et al., 1978; Fu et al., 2013) and more recently to multi-cellular organisms (Gogarten and Townsend, 2005; Iyer et al., 2011) suggesting a more wide-spread presence of 6mA in eukaryotes. Thus, 6mA was recently detected in several metazoan genomes including, *Drosophila* (Zhang et al., 2015), *Caenorhabditis elegans* (Greer et al., 2015a), and, controversially, in vertebrate genomes such as mouse (kidney cells, (Koziol et al., 2016); embryonic stem cells, (Wu et al., 2016a), *Danio rerio* and *Sus scrofa* (Liu et al., 2016).

### 1.1.4. Writers and erasers of deoxyadenosine methylation

6mA is generated through the transfer of a methyl group from S-adenosyl-L-methionine to the exocyclic NH2 in position 6 of the purine ring ($N^6$) by DNA methyltransferases (Gold et al., 1963). First observation of this was reported in *Escherichia coli* cell extracts where the DNA adenine methylases such as Dam and MunI thought to selectively methylate target adenine bases (Gold et al., 1963; Srinivasan and Borek, 1964; Hurwitz et al., 1965; Nikolskaya et al., 1981; Blow et al., 2016; Harris and Goldman, 2020). In similar manner, various eukaryotic genomes have been shown to contain putative N6-adenosine-methyltransferases with an MT-A70 domain. These include DNA 6mA methyltransferase complex in the ciliate *Oxytricha*, termed MTA1c (Beh et al., 2019; Harris and Goldman, 2020), DNA N6 adenine methyltransferase 1, DAMT-1, in *C. elegans* (Greer et al., 2015a), Ime4-like methyltransferase in *Drosophila* (Zhang et al., 2015), METTL4 in human mitochondria (Hao et al., 2020) and N-6 adenine-specific DNA methyltransferase 1 (N6MT1)

in mammals (Kweon et al., 2019). Despite the identification of several candidate 6mA demethylases, the molecular mechanisms of 6mA demethylation in multicellular eukaryotes remain unclear. For example, even though a *Drosophila* DAMD, a homolog of ten-eleven translocase (TET) was recently proposed to demethylate 6mA (Zhang et al., 2015), no evidence for TET-mediated 6mA demethylation reported in *Drosophila* (Iyer et al., 2016; O'Brown and Greer, 2016), raising questions regarding the efficiency of DAMD mediated 6mA demethylation *in vivo*. Nevertheless, other Fe (II)- and α-ketoglutarate-dependent dioxygenase members such as *C. elegans* NMAD-1 as well as mammalian ALKBH1 and ALKBH4 have also been proposed as potential 6mA demethylases (Greer et al., 2015a; Kweon et al., 2019; Xiao et al., 2018).

### 1.1.5. Biological consequences of deoxyadenosine methylation in eukaryotes

Through the identification and manipulation of 6mA reader proteins, 6mA involvement in DNA replication (Lu et al., 1994), restriction-modification system (Iyer et al., 2011), and relaxation of double stranded DNA (Cheng et al., 1985) has been shown for prokaryotes. Apart from its recent detection in multicellular eukaryotes (Fu et al., 2015; Greer et al., 2015b; Zhang et al., 2015), very little is known about the function of eukaryotic 6mA. Nonetheless, in *C. elegans* 6mA was found to trans-generationally accumulate following the depletion of *spr*-5, a histone H3 lysine 4 dimethyl (H3K4me2) demethylase, or after the deletion of *nmad-1*, a DNA m6A demethylase, suggesting association of 6mA with active transcription (Greer et al., 2015a). Consistent with 6mA association with transcriptional activity, in *Drosophila*, elevated levels of 6mA were observed in gene bodies of actively transcribed transposons (Zhang et al., 2015). In contrast, in higher eukaryotes, where the presence of 6mA is still controversial (Douvlataniotis et al., 2020; Lentini et al., 2018), 6mA has been reported to correlate with gene repression by preserving polycomb mediating transcriptional silencing (Kweon et al., 2019).

### 1.1.6. Contradictory reports of deoxyadenosine methylation in mammals.

Even though the presence of 6mA in lower eukaryotes has been widely confirmed, there is an ongoing controversy regarding reports of 6mA detection in mammalian genomes. Indeed, the controversy started soon after the first reported 6mA detection in mammals in studies of bull and human sperm DNA in the 1960s (Unger and Venner, 1966). Despite painstaking efforts, other groups were not able to reproduce this finding (Vanyushin et al., 1970), and as a result the findings were dismissed with 6mA presence interpreted as a result of potential contamination of the mammalian samples with prokaryotic DNA. Using a dot-blot method, Achwal and colleagues reported the detection of 6mA in several species including DNA preparations from human placental tissue (Achwal et al., 1983). Subsequently, using a combination of dot-blot and DNA immunoprecipitation, 6mA was reportedly detected in mouse (Wu et al., 2016a; Yao et al., 2017) and human (Huang et al., 2015) cells. Attempting to settle the controversy, recently, the Carell lab performed analysis of mouse embryonic stem cells and tissues using a novel ultrasensitive UHPLC-MS method but found no 6mA or 4mC in these samples (Schiffers et al., 2017). Corroborating these findings, combined bioinformatics and experimental analysis highlighted potential sources of contamination responsible for apparent presence of 6mA in mammalian samples (O'Brown and Greer, 2016; Douvlataniotis et al., 2020). As the inherently low amounts of 6mA in eukaryotic genomes further fuel this controversy (Ratel et al., 2006), studies of 6mA in mammalian cells requires not only a carefully designed and properly controlled analysis but also highly sensitive techniques.

## 1.2. N6-methyladenosine (m6A) in RNA

### 1.2.1. Methylation of RNA adenosine is dynamic

Unlike its DNA counterpart, the presence of N6-methyladenosine (m6A) in eukaryotic RNAs has been well documented since the 1970s (Desrosiers et al., 1974; Schibler et al., 1977).

Indeed, amongst the hundreds of known RNA modifications (Dal Magro et al., 2018; Machnicka et al., 2013), m6A is the most abundant internal modification of mRNAs and long noncoding RNAs (lncRNAs) of mammalian cells (Cantara et al., 2011; Sommer et al., 1978). Interestingly, m6A presence on mRNA has been verified in species from all three domains of life, including viruses (Krug et al., 1976), bacteria (Deng et al., 2015), plants (Zhong et al., 2008), yeast (Clancy et al., 2002), and mammalian cells (Desrosiers et al., 1974; Perry and Kelley, 1975). In addition to decorating messenger RNAs (Desrosiers et al., 1974; Carroll et al., 1990), adenine methylation occurs in other RNA classes, including ribosomal RNAs (Maden, 1986; Sergiev et al., 2008), transfer RNAs (Björk et al., 1987), small nuclear RNAs (Shimba et al., 1995), microRNAs (Alarcón et al., 2015) and long noncoding intergenic RNAs (Barbieri and Kouzarides, 2020; Fazi and Fatica, 2019). Despite its widespread presence across transcripts, on average, only 3 methylated adenosines are detected per messenger RNAs (Desrosiers et al., 1974; Perry and Kelley, 1975), most frequently in the Pu[G > A]$m^6$AC[U > A > C] context where Pu denotes purine (Rottman et al., 1994; Shimba et al., 1995). Using a combination of anti-m6A specific antibody and high throughput next generation sequencing approaches, recently, the transcriptome-wide distribution maps of m6A across several eukaryotic species have been generated for a number of species including human (Dominissini et al., 2012), mouse (Batista et al., 2014) and yeast (Schwartz et al., 2013). These maps rely on affinity capture methods where the extracted mRNA is fragmented, m6A containing transcripts are pulled down and sequenced following reverse transcription and library generation from the pulldown and input control (Dominissini et al., 2012; Meyer et al., 2012). These studies revealed transcriptome-wide m6A enrichment near stop codons, within 3' UTR regions with about 70% of m6A-marked transcripts conserved between human are mouse, suggesting a conservation of function in these organisms (Batista et al., 2014; Dominissini et al., 2012; Meyer et al., 2012; Schwartz et al., 2013). Although

recent introduction of base-resolution m6A mapping methods (Ke et al., 2015; Linder et al., 2015) allowed a more precise mapping of m6A profiles transcriptome-wide (Koh et al., 2019), it remains largely unclear how the region specific deposition of m6A along the transcript (e.g. enrichment in 3'UTR as opposed to exons) is achieved. The discovery of m6A writers, erasers and readers (Figure 1.2.1) might shed light on the precise molecular function of m6A (He, 2010).

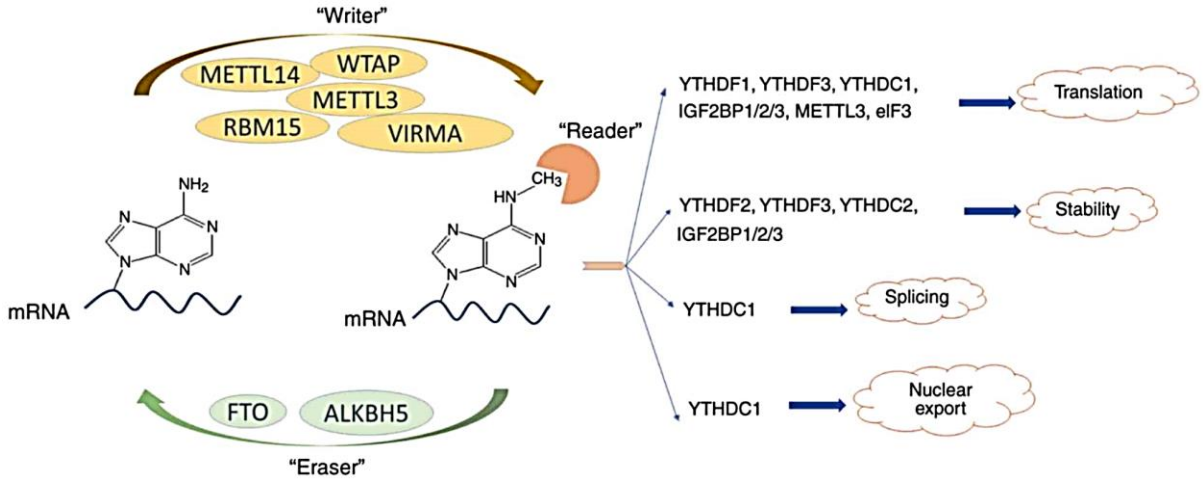



**Figure 1.2.1. The dynamic and reversible nature of m6A RNA methylation.** This is mediated by a group of writer and eraser proteins while the presence of the mark is interpreted by a group of reader proteins (adapted from (Zhu et al., 2019).

### 1.2.2. Writers of m6A

The presence of chemical modifications on RNA bases has been known for more than 4 decades (Desrosiers et al., 1974; Wei et al., 1975), the prevailing view was that RNA modifications are static and bear no important functional roles in RNA metabolism (Yue et al., 2015). This view was largely based on the fact that unlike genomic DNA, RNAs have short half-lives (Rabani et al., 2011; Schwanhäusser et al., 2011). Moreover, the m6A demethylases were until recently unknown (He, 2010). Almost two decades after the discovery of m6A on messenger RNAs, Bokar and colleagues conclusively showed that a 70-KD protein termed MT-A70, isolated from mammalian cell extracts as a part of a 200-KD

complex, exhibits methyltransferase activity based on its ability to bind S-Adenosyl methionine (SAM) and install the methyl group on target RNA adenosines (Bokar et al., 1994, 1997). The ability of this methyltransferase to generate m6A was further demonstrated both *in vitro* and *in vivo* (Narayan and Rottman, 1988). Indeed, METTL3 or MT-A70 depletion was shown to restrict levels of m6A methylation (Bokar, 2005). The m6A methyltransferase enzymes are conserved from yeast to humans; in *Saccharomyces cerevisiae* METTL3 ortholog IME4 (inducer of meiosis 4) has been shown to generate m6A during meiosis (Agarwala et al., 2012; Schwartz et al., 2013), in *Drosophila melanogaster*, the METTL3 ortholog, IME4 was shown to be essential for m6A methyltransferase activity as well as normal development (Hongay and Orr-Weaver, 2011) and in various plants that include *Arabidopsis*, maize, and rice, the METTL3 orthologue, MTA, is ubiquitously expressed and exhibits m6A methyltransferase activity (Haugland and Cline, 1980; Zhong et al., 2008). In eukaryotic systems, including mammals, the Wilms' tumor 1-associating protein (WTAP) (Little et al., 2000) has been shown to act as a co-factor in the METTL3-METTL14 methyltransferase complex and, thus, to ensure efficient methylation of target adenosines (Agarwala et al., 2012; Ping et al., 2014). Recently, through the identification of WTAP-dependent and independent m6A deposition sites, WTAP association with the METTL3/METTL14 complex is thought to preferentially determine 5' and internal mRNA sequences (Schwartz et al., 2013; Sun et al., 2015), thereby potentially contributing to regional specificity of m6A deposition along transcripts. In addition to WTAP, in recent years, through proteomic and functional analysis (Horiuchi et al., 2013; Schwartz et al., 2013), additional members of the METTL3/METTL14 multiprotein complex have been identified in mammals (Ping et al., 2014). These include, KIAA1429, a mammalian homolog of the *Drosophila* Virilizer protein (Schwartz et al., 2013), the RNA binding motif protein 15 (RBM15) and its paralogue RBM15B targeting METTL3 to RNA (Lence et al., 2016; Patil et

al., 2016). While it is becoming increasingly clear that METTL3/METTL14 constitute the core component of the methyltransferase complex, more recently, additional methyltransferases have been identified (Pendleton et al., 2017) which have been shown to methylate long noncoding RNAs and intron sequences in CAG contexts as opposed to the canonical DRACH motif (Nance et al., 2020; Wang et al., 2015; Warda et al., 2017). What is becoming increasingly obvious is that methyltransferases co-opt with other proteins to not only identify their target RNAs (Ping et al., 2014) but also discriminate between different adenosine contexts for methylation.

### 1.2.3. Erasers of m6A

In a landmark study, the He group demonstrated that the Fat mass and obesity-associated (FTO) protein isolated from nuclear extracts has m6A demethylase activity (Jia et al., 2011). FTO is a member of the non-heme Fe(II)- and α-KG-dependent AlkB family of dioxygenases (Gerken et al., 2007; Kurowski et al., 2003) that has been shown to oxidatively demethylate products of alkylation damage present in both single stranded RNA (3-methyluracil (Jia et al., 2008) and single stranded DNA (3-methylthymidine, Aas et al., 2003; Gerken et al., 2007), but the He group was the first to demonstrate that m6A may serve as a substrate for FTO mediated demethylation (Jia et al., 2011). Akin to oxidative demethylation of methylated DNA cytosines by ten-eleven (TET) proteins that convert 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5fC) (Tahiliani et al., 2009; Ito et al., 2010, 2011), FTO has been shown to oxidatively convert m6A to N6-hydroxymethyladenosine and N6-formyladenosine (Fu et al., 2013). Due to the inherent instability of FTO-mediated m6A demethylation intermediates (Yue et al., 2015), it is unknown whether they have any separate biological functions. More recently, FTO has been shown to exhibit higher demethylation activity towards $N^6$,2-O-dimethyladenosine ($m^6Am$) compared to m6A (Mauer et al., 2017). In addition to FTO, the He group also identified

ALKBH5 as another m6A demethylase (Zheng et al., 2013). Despite the ability of both FTO and ALKBH5 to demethylate m6A, there is an appreciable difference in their subcellular localisation and expression pattern in a cell- and tissue-type-specific manner (Meyer and Jaffrey, 2014). Whereas FTO is detectable in both nucleus and, albeit weaker levels, in the cytoplasm (Hess et al., 2013), ALKBH5 expression is thought to be restricted to the nucleus (Zheng et al., 2013). Moreover, FTO is mainly expressed in the brain, whereas ALKBH5 protein is predominantly detected in testis (Zheng et al., 2013). Such distinct localisation and cell-type-specific distribution of FTO and ALKBH5 demethylases may indicate difference in their target specificity. Nevertheless, the discovery of m6A demethylases represent a paradigm shift in our understanding of the dynamic and reversible nature of RNA modifications in eukaryotic transcriptomes, and, thus, it ushers a new era of epitranscriptomic studies.

### 1.2.4. Readers of m6A

In addition to m6A demethylases, multiple m6A dependent RNA binding proteins have been captured in pull-down assays using m6A modified RNA probes and their identity uncovered by mass spectrometry analysis (Dominissini et al., 2012; Schwartz et al., 2013). YTH (YT521-B homology)-domain containing protein family proteins, YTHDF1, YTHDF2 and YTHDF3 were amongst the first identified for their high affinity to m6A modified RNAs *in vitro* (Dominissini et al., 2012; Schwartz et al., 2013), and later functionally validated using knockdown experiments in mammalian cells (Wang et al., 2014). YTH-domain is a conserved RNA biding domain found in over 100 eukaryotic proteins (Zhang et al., 2010) that executes m6A binding affinity via a conserved aromatic cage (Liao et al., 2018; Xu et al., 2015). In addition to YTHDF1-3 proteins, there are two more YTH-domain containing proteins, YTHDC1 and YTHDC2 (Zhang et al., 2010). Structural analysis show that YTHDC1 preferentially interacts with m6A in GGm6AC sequence contexts via an aromatic

cage in its YTH-domain (Xu et al., 2015), in a similar fashion to royal family protein recognition of methyllysine (Yap and Zhou, 2011). In contrast to other human YTH-domain containing proteins, YTHDC2, beside its YTH-domain, also contains a number of other domains: a domain of unknown function (DUF1065), an ankryin repeats (ANK) domain and two putative helicase domains (Figure 1.2.4) (Liao et al., 2018) that may determine various biological consequences of YTHDC2 interaction with m6A (Kretschmer et al., 2018).

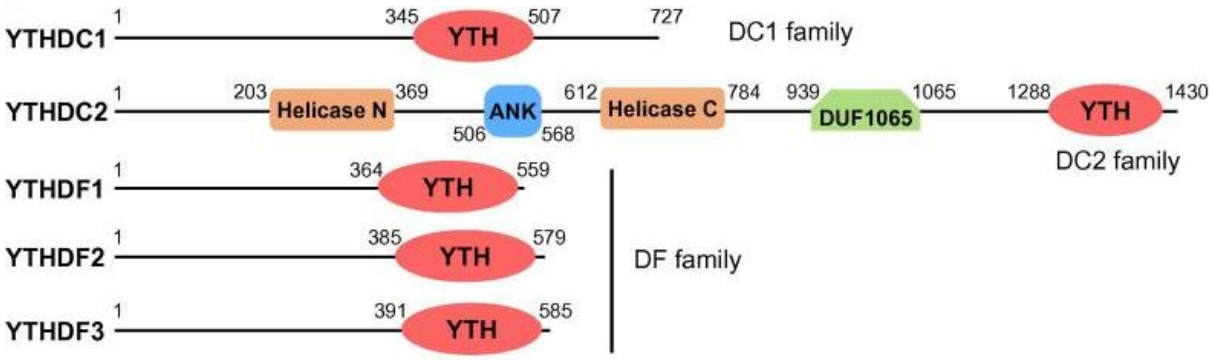



**Figure 1.2.4. Domain architecture of human YTH domain-containing proteins**: Schematic showing the domains of YTHDC1 (UniProt ID: Q96MU7), YTHDC2 (UniProt ID: Q9H6S0), YTHDF1 (UniProt ID: Q9BYJ9), YTHDF2 (UniProt ID: Q9Y5A9), and YTHDF3 (UniProt ID: Q7Z739), adapted with modifications from (Liao et al., 2018).

Beside YTH-domain containing proteins, m6A is recognised by other RNA binding proteins (Zhao et al., 2020), these include heterogeneous nuclear ribonucleoproteins HNRNPA2B1 which binds m6A in microRNAs (Alarcón *et al*., 2015), HNRNPC and HNRNPG (Zarnack et al., 2013) that thought to bind m6A modified RNAs (Liu et al., 2017; Spitale et al., 2015) as a result of m6A-induced structural changes in RNAs (Liu et al., 2017; Spitale et al., 2015) Moreover, insulin-like growth factor 2 mRNA-binding proteins 1, 2 and 3 have been shown to bind m6A marked mRNAs through their K homology domain (Huang et al., 2018; Müller et al., 2019). More recently, proline-rich and coiled-coil-containing protein 2A (PRRC2A) has been reported to interact with m6A bearing transcripts via GGACU sequence motif (Wu

et al., 2019). Together, these m6A binding protein continue to reveal the diverse functional roles of m6A.

### 1.3. m6A deposition impacts multiple steps of RNA biogenesis

#### 1.3.1. Role of m6A in RNA Stability and degradation

As early as in the 1970s, using [$^{3}$H] or $H_3{}^{32}PO_4$ adenosine labelled HeLa cells, Darnell and colleagues showed that m6A presence leads to faster turnover of transcripts (Sommer et al., 1978; Ke et al., 2015). However, the mechanism remained largely unknown until the He lab recently demonstrated that YTHDF2 proteins specifically interact with m6A located in predominantly GGAC sequence motifs on both messenger and noncoding RNAs to target them to RNA degradation sites such as P-bodies (Wang et al., 2014). Importantly, YTHDF2 knockdown showed extended half-lives of m6A marked RNA (Wang et al., 2014), thereby unveiling a critical role for YTHDF2 in regulation the stability of m6A-modified RNAs. Subsequently, YTHDF2 has been shown to mediate degradation of m6A marked RNAs by recruiting the CCR4-NOT deadenylase complex to corresponding RNAs for decay (Du et al., 2016). Alternatively, YTHDF2 has also been shown to mediate degradation of m6A marked RNAs via its interaction with the human heat-responsive protein 12 (HRSP12) and subsequent cleavage of the RNA with the RNase P/MRP endoribonuclease complex (Park et al., 2019). In a recently published work, the Jaffrey lab presented data that suggests all YTHDF proteins (YTHDF1,2 &3) have a unified role in degradation of m6A marked transcripts (Zaccara and Jaffrey, 2020). Thus, their report re-emphasises a fundamental role for m6A in regulating RNA stability. Moreover, although this finding is in agreement with the functional compensation between YTHDF proteins, recently reported by the Hannah group (Lasman et al., 2020), what underlies the dramatic phenotypic differences between the individual knockouts of each of the YTHDF proteins remain largely a mystery. For example, while YTHDF2 KO mice exhibit pre-weaning embryonic lethality (Li et al., 2018), neither

YTHDF1 nor YTHDF3 KO mice seem to have such a drastic phenotype (Lasman et al., 2020; Shi et al., 2018; Zhang et al., 2019). This raises an interesting question of whether the functional redundancy is temporally and spatially restricted. Nevertheless, the supreme role of YTHDF2 in mediating the degradation of m6A modified RNAs remain unchallenged. Conversely, m6A has also been shown to recruit HuR proteins that result in stabilisation of a select m6A containing RNAs (Wang et al., 2015; Huang et al., 2018). This reader-dependent function of m6A was also observed in METTL3 depletion studies, where increased stability of m6A depleted transcripts has been detected (Wang et al., 2014; Ke et al., 2017).

### 1.3.2. The role of m6A in regulating RNA splicing.

The m6A presence on pre-mRNA has been shown to regulate alternative splicing both in human (Zhao et al., 2014) and flies (Haussmann et al., 2016). The m6A nuclear reader, HNRNPA2B1 has also recently been shown to modulate the splicing of micro RNAs in an m6A-dependent manner (Alarcón et al., 2015). More recently, another nuclear m6A reader, YTHDC1 has been shown to promote alternative splicing of oocyte transcripts as well as regulate mRNAs' polyadenylation (Kasowitz et al., 2018). Taken together, these studies reveal an additional layer of RNA processing control mediated by m6A and its reader proteins (Zhu et al., 2020).

### 1.3.3. Role of m6A in export of nuclear RNA

In one of the earliest indications for the role of m6A in promoting RNA nuclear export, cycloleucine, a potent inhibitor of methionine transferase, was shown to suppress nuclear export of mRNA (Finkel and Groner, 1983). Subsequently, direct inhibition of METTL3 has been shown to slow the nuclear export of clock gene transcripts per2 and Arntl and, as a result, a longer circadian period has been observed (Fustin et al., 2013). Recently, the m6A reader, YTHDC1 has also been implicated in promotion of nuclear export of m6A-marked RNAs as its depletion led to nuclear retention of m6A modified mRNAs (Roundtree et al.,

2017). Extending these findings, Wilson and colleagues confirmed the role of YTHDC1 in nuclear export and further demonstrated that the m6A machinery recruit the multi-subunit mRNA export machinery, TREX, for nuclear export of m6A marked transcripts (Lesbirel et al., 2018). More recently, m6A presence on select noncoding RNAs has been shown to accelerate their export (Zuckerman et al., 2020). In like manner, inhibition of the m6A demethylase, ALKBH5 has been shown to promote the increased nuclear export of m6A marked RNAs (Zheng et al., 2013). Collectively, these studies reveal an important role of m6A in promoting nuclear to cytoplasmic export of target RNAs.

### 1.3.4. The role of m6A in the regulation of RNA translation.

The first indication of m6A function in the regulation of RNA translation was presented by He and colleagues (Wang et al., 2015). They show that YTHDF1 interacts with m6A-modified RNA and promote ribosome loading of m6A marked RNAs (Wang et al., 2015). Although Wang *et al*did not reportchanges in the ratio of m6A modified and nonmodified RNAfollowing YTHDF1 knockdown (Wang et al., 2015), more recent reports suggest that all YTHDF proteins (YTHDF1/2/3) exhibit functional redundancy in regulating mRNA stability but not translation (Zaccara and Jaffrey, 2020). This latest study by the Jaffrey group seems to challenge earlier described consequences of m6A in promoting translation of m6A marked RNA via YTHDF1 (Wang et al., 2015) but also fails to explain the obvious disparities in the reported phenotypes of YTHDF3 KO (mild) compared to YTHDF2 (embryonic lethal )(Li et al., 2018; Shi et al., 2018).

## 1.4. Cellular and physiological functions of m6A modified RNA

### 1.4.1. The role of m6A in meiosis

First insights into the physiological and functional role of m6A came from studies in *Saccharomyces cerevisiae* that showed that m6A appears on target mRNA exclusively during meiosis (Shah and Clancy, 1992; Clancy et al., 2002). Deposition of m6A on mRNA during

meiosis is mediated by IME4 (inducer of meiosis 4), an orthologue of mammalian m6A writer, METTL3 (Agarwala et al., 2012). Although IME4-depleted yeasts survive, the absence of IME4 was found to delay entry into meiotic division that is required for proper sporulation (Shah and Clancy, 1992; Agarwala et al., 2012; Schwartz et al., 2013). In contrast, complete loss of IME4 in *Drosophila melanogaster* has been shown to result in lethality (Hongay and Orr-Weaver, 2011). However, hypomorphic *Drosophila melanogaster* mutants of IME4 have been shown to exhibit aberrant notch signalling that in turn lead to defective oogenesis (Hongay and Orr-Weaver, 2011; Kan et al., 2017). Recently, these meiotic functions of m6A has also been observed in mammals. For example, mice with testis specific METTL3 deficiency exhibit impaired meiosis initiation and sterility (Xu et al., 2017). Conversely, deletion of the m6A demethylase, ALKBH5, has been shown to result in defective spermatogenesis, due, in part, to accumulation of m6A marked transcripts linked with cell-death in spermatocytes undergoing meiosis (Zheng et al., 2013). Recently, YTHDF2 has been shown to mediate the selective degradation of m6A marked mRNAs during oocyte development with its deletion leading to defective oogenesis and subsequent infertility (Ivanova et al., 2017). More recently, the involvement of another m6A reader protein, YTHDC2, in meiotic initiation and progression of mouse germ cells was reported as its depletion was found to impair proper meiosis (Zeng et al., 2020). It is still largely unclear to what extent the reported meiotic functions of m6A reader proteins is a consequence of their interaction with m6A.

### 1.4.2. The role of m6A during stress response

Motivated by the dynamic nature of m6A modification (Jia et al., 2011), several studies uncovered a role for m6A in cellular and organismal stress response. The m6A reader protein, YTHDF2 was found to protect m6A deposited on 5’ UTR of stress induced transcripts by translocating to the nucleus during heat shock stress (Zhou et al., 2015). The reported

YTHDF2 translocation to the nucleus was shown to be required for protecting heat-induced m6A modified transcripts from demethylation by FTO (Zhou et al., 2015), such that they can participate in cap-independent translation via interaction with eIF3 (Meyer et al., 2012; Meyer and Jaffrey, 2014). In a recent paper, Zhou *et al*. extended these findings by revealing the involvement of m6A in amino-acid starvation-induced stress response (Zhou et al., 2015). Direct perturbation of RNA methylation levels by deletion of the writer, METTL3 or methylation eraser, FTO, in adult mice neurons compromised fear-induced stress response (Engel et al., 2018). Similarly, upon ultra-violet mediated DNA damage, METTL3 mediated m6A deposition was observed at the site of damage (Xiang et al., 2017). Collectively, these studies further establish m6A as a dynamic modification that can provide cells and organisms with a stimulus dependent response.

### 1.4.3. The role of m6A in cellular differentiation and development

Initial hints for m6A physiological role came from studies in *yeast* that showed knockdown of IME4, the yeast METTL3 ortholog, resulted in sporulation defects (Shah and Clancy, 1992; Agarwala et al., 2012; Schwartz et al., 2013). In multicellular eukaryotes, complete loss of the m6A methyltransferase has been shown to often result in lethality. For example, in *Drosophila melanogaster*, complete loss of IME4 is lethal (Hongay and Orr-Weaver, 2011) while, hypomorphic mutants of IME4 have been shown to exhibit aberrant notch signalling that in turn lead to defective oogenesis (Hongay and Orr-Weaver, 2011; Kan et al., 2017). In zebrafish, disruption of mRNA clearance through targeting m6A reader YTHDF2 has been shown to result in defective zygotic genomic activation (Zhao et al., 2017). Similar effect of YTHDF2 removal has also been documented in mice, where mice oocytes lacking YTHDF2 failed to undergo proper maternal to zygotic genome activation (Ivanova et al., 2017) largely due to defective oocyte competence (Lasman et al., 2020). Direct disruption of m6A patterns through complete METTL3 deletion in mice proved to be challenging due to its embryonic

lethal phenotype (Geula et al., 2015). However, mouse and human METTL3 knockout ES cells are viable but fail to differentiate properly (Dominissini et al., 2012; Geula et al., 2015), largely due to prolonged expression of pluripotency factors, NANOG and SOX2 (Batista et al., 2014). Tellingly, METTL3 KO mouse embryos die between embryonic day 3.6 to 6.5 which is known to coincide with onset of differentiation (Batista et al., 2014). In a similar manner to METTL3 deletion, METTL14 KO mice also failed to properly transition through epiblast stage, further confirming the role of m6A in regulating cellular differentiation and development (Meng et al., 2019). Investigation of SMAD2/3 interactome in human pluripotent stem cells by Bertero *et al* revealed that SMAD2/3 interacts with METTL3/14 to guide m6A deposition on TGFβ/activin signalling pathway transcripts linked with pluripotency (Bertero et al., 2018). In further support for m6A role in promoting cellular differentiation, unlike their wild type counterparts, METTL3 KO hematopoietic stem cells failed to properly differentiate upon stimulation due to the failure of expression of MYC mRNA that is m6A-modified in wild type cells (Lee et al., 2019). The role of m6A in regulating differentiation and development also extends to the nervous system (Hsu et al., 2017) where the presence of m6A-enriched transcripts is the highest amongst different tissues (Meyer et al., 2012; Chang et al., 2017). Disruption of m6A deposition via METTL14 depletion has been shown to impair axonal regeneration in adult mouse dorsal root ganglion (Weng et al., 2018). Similarly, disruption of m6A patterns via METTL3 depletion in mouse embryonic brain tissues has been shown to result in delayed cell-cycle of radial glia that, in turn, extended cortical neurogenesis (Yoon et al., 2017). Since such decreased proliferation rate was also observed upon FTO depletion (Li et al., 2017), it may suggest that a tight balance between the rate of m6A deposition to and removal from different transcripts is key to ensure proper differentiation and development of the brain (Li et al., 2017).

### 1.4.4. The role of m6A in disease

In addition to its role in development, by virtue of its reversible nature and its major role in almost all steps of RNA metabolism, it is perhaps unsurprising that links between m6A mis-regulation and multiple diseases have been uncovered (Barbieri and Kouzarides, 2020). High levels of the m6A methyltransferase, METTL3 was found in acute myeloid leukaemia cells compared to healthy hematopoietic stem/progenitor cells (Vu et al., 2017). Depletion of METTL3 using short heterogeneous RNA (shRNA) treatment was found to promote differentiation, death of leukemic cells or delayed leukaemia progression (Vu et al., 2017), suggesting the potential therapeutic target of METTL3 in myeloid leukaemia (Vu et al., 2017). The Kouzarides lab confirmed these findings and further identified a select group of gene promoters whose binding by METTL3 is found to be essential to promote translation of their transcripts and subsequent myeloid leukaemia progression (Barbieri et al., 2017). Soon after these reports and in similar manner to METTL3 (Vu et al., 2017), METTL14 has been shown to promote self-renewal and progression of acute myeloid leukaemia as its suppression promoted terminal myeloid differentiation (Weng et al., 2018). In all three studies, the role of METTL3/METTL14 deposited m6A in promoting translation of oncogenic transcripts is consistently highlighted (Barbieri et al., 2017; Vu et al., 2017; Weng et al., 2018). METTL3 over-expression has also been implicated in many cancer types. In lung cancer, METTL3 over-expression thought to promote translation of oncogenic transcripts through association with the translation initiation machinery (Lin et al., 2016). Similarly, in liver cancer, METTL3 over-expression was found to promote the translation of oncogenic transcripts that are critical for epithelial to mesenchymal transition (Lin et al., 2019). In addition to m6A methyltransferases, there is a growing appreciation for the role of m6A demethylases in diseases. For example, a consistently increased expression of FTO has been reported in many cases of human cancers (S. Yang et al., 2019; Barbieri and Kouzarides, 2020). In human

melanoma cells, FTO overexpression thought to promote melanoma cell growth by exposing key transcripts involved in tumour suppression to YTHDF2 decay, these include *PD-1*, C-X-C motif chemokine receptor 4, and SRY-Box transcription factor 10 (S. Yang et al., 2019). Similarly, FTO and YTHDF1 has also been shown to promote tumour cell growth and differentiation in acute myeloid leukaemia and melanoma respectively (Li et al., 2017). The association of FTO over-expression with tumorigenicity of cancer cells has also been reported in breast cancer cells (Niu et al., 2019). In a similar manner to FTO, ALKBH5 over-expression has also been shown to promote proliferation, migration, and invasion of cancer cells (Barbieri and Kouzarides, 2020). For example, ALKBH5 over-expression has been shown to promote the stability of plasmacytoma variant translocation 1, an oncogenic long noncoding RNA, in osteosarcoma cells (Chen et al., 2020). Although mounting evidence demonstrate the critical roles of m6A writer and eraser proteins in cancer (Barbieri and Kouzarides, 2020), it remains largely unclear to what extent the reported roles of these proteins in cancer are dependent on their ability to bind m6A. Nonetheless, perturbation of m6A reader proteins has also been implicated in the pathogenesis of many cancers (Barbieri and Kouzarides, 2020). For example, YTHDF1 over-expression was found to promote ovarian cancer metastasis by increasing the translation of targets transcripts in an m6A-dependent fashion (T. Liu et al., 2020), a feature that has been document in other cancers (Barbieri and Kouzarides, 2020). YTHDF2 misexpression has also been implicated in initiation and/or progression of multiple cancers, ranging from acute myeloid leukemia; where YTHDF2 was found to mediate the decay of m6A modified pro-apoptotic tumour necrosis factor signalling pathway transcripts (Paris et al., 2019). In prostate cancer, YTHDF2 over-expression was found to promote proliferation of prostate cancer cell lines, and its inhibition suppressed proliferation and migration of prostate cancer cells (T. Liu et al., 2020), whereas in hepatocellular carcinoma YTHDF2 has been suggested to promote

metastasis by compromising the immune system through the decay of m6A-containing interleukin 11 and serpin family E member 2 transcripts (Hou et al., 2019). Intriguingly, a common theme is emerging where changes in m6A levels (directly or indirectly) seem to impact the rate of cell cycle both in healthy (cell resisting differentiation, (Lasman et al., 2020) and in cancer cells (promoting tumorigenicity, (Barbieri and Kouzarides, 2020). Despite this, very little is known about the role of m6A in the cell-cycle regulation.

### 1.5. RNA:DNA hybrids (R-loops).

#### 1.5.1. Definitions and potential origins

Existence of a rudimentary form of RNA:DNA hybrids is thought to predate the very emergence of biological systems (Bhowmik and Krishnamurthy, 2019; J. Xu et al., 2020). Therefore, it is perhaps unsurprising that RNA:DNA hybrids are not only present in all three domains of life (Carles-Kinch and Kreuzer, 1997; Hyjek et al., 2019; Yu et al., 2003) but perform essential functions in numerous biological processes including such basic ones as the replication of DNA lagging strand by synthesis of Okazaki fragments (Sugino et al., 1972), extension and protection of telomeres (Greider and Blackburn, 1989), erroneous incorporation of ribonucleotides during DNA leading strand synthesis (Lujan et al., 2013; Nick McElhinny et al., 2010) and within the active site of RNA polymerase II during transcription initiation (Nudler et al., 1997). All the aforementioned RNA:DNA hybrids are typically made up of less than 11 base pairs (Costantino and Koshland, 2018). However, during the transcription of nascent RNA, the newly synthesised RNA can, on exit from the active site of RNA polymerase, invade the negatively supercoiled and therefore partially unwound double stranded DNA (Liu and Wang, 1987), this permits the RNA to anneal to its complementary template DNA strand, generating a longer stretch of RNA:DNA hybrid and displace the single stranded DNA, in sum, giving rise to a three-stranded nucleic acid structure, known as R-loop (Thomas et al., 1976). Although structural characterisation of

such RNA:DNA hybrids was reported since 1970s (White and Hogness, 1977), evidence for an *in vivo* formation of RNA:DNA hybrids was first reported in 1995 by the Crouch lab in *Escherichia coli* study (Drolet et al., 1995). Subsequent studies uncovered the presence of R-loops in *Saccharomyces cerevisiae* (Wahba et al., 2011), *Caenorhabditis elegans* (Castellano-Pozo et al., 2012), *Drosophila Melanogaster* (Alecki et al., 2020), zebrafish (Sorrells et al., 2018), mice (Sanz et al., 2016a), humans (Ginno et al., 2012, 2013) as well as plants (W. Xu et al., 2020). Despite the prevalence of R-loops, the precise mechanism of how R-loops formation occurs *in vivo* remain largely unclear (Skourti-Stathaki and Proudfoot, 2014). Nonetheless, two models have been proposed, an "extended RNA/DNA hybrid" and "thread back" model (Westover et al., 2004). The "extended RNA:DNA hybrid" model suggests that R-loops arise because of an extension to the RNA:DNA hybrid formed within the active site of RNA polymerase (Westover et al., 2004). In contrast, the "thread-back" model suggests that on release from the active site of an RNA polymerase, the nascent RNA "threads-back" and re-anneals to its complementary DNA strand by invading the double stranded DNA and displacement of a single stranded DNA (Westover et al., 2004). However, the former model is not supported by crystallographic structure of RNA Pol II that uncovered the existence of two distinct exit channels for the nascent RNA and the template DNA (Westover et al., 2004). Instead, the thread-back model is more widely accepted (Aguilera and García-Muse, 2012; Skourti-Stathaki and Proudfoot, 2014) due to observations that formation of RNA:DNA hybrids occurs co-transcriptionally e.g. in response to defective transcriptional elongation, splicing or termination (Costantino and Koshland, 2015). Such transcriptional defects thought to promote formation of R-loops by providing nascent RNA exiting the RNA polymerase an extended opportunity to re-anneal to the double stranded DNA, and therefore generating an R-loop (Costantino and Koshland, 2015). In addition to defective RNA processing, several structural features in the DNA have been shown to promote RNA:DNA

hybrid formation. These include, transcription from a genomic region rich in cytosine clusters or a G-rich transcript (Yu et al., 2003; Ginno et al., 2013; Chan et al., 2014), the presence of specific DNA secondary structures, G-quadruplexes (Sundquist and Klug, 1989), on the non-template strand (De Magis et al., 2019), negative supercoiling of DNA template (Drolet, 2006; Roy et al., 2010) and the presence of nicks on the non-template DNA strand (Roy et al., 2010). Although the thread-back model necessitates co-transcriptional formation of RNA:DNA hybrids (Huertas and Aguilera, 2003; Aguilera and García-Muse, 2012), recent evidence challenges the universality of this mechanism. For example, R-loops can also form in response to abnormal processing of post-transcriptional events, such as RNA export (Huertas and Aguilera, 2003) and defective degradation of RNA (Costantino and Koshland, 2015). In addition, RNA:DNA hybrids have also been shown to form *in trans*, away from the initial site of transcription, evidence for this was first described in *yeast* where induction of indigenous transcript was shown to form an RNA:DNA hybrid with an artificial yeast chromosome (Wahba et al., 2013).

### 1.5.2. Consequences of RNA:DNA hybrid formation

As the thermodynamic stability of RNA:DNA hybrids is higher than that of double stranded DNA (Roberts and Crothers, 1992), regardless of how they form, the mere presence of RNA:DNA hybrids in the genome have been found to influence a number of important biological processes (Santos-Pereira and Aguilera, 2015; Niehrs and Luke, 2020), contributing to both physiological and pathological consequences (Aguilera and García-Muse, 2012; Crossley et al., 2019).

### 1.5.2. The role of RNA:DNA hybrids in DNA replication

The earliest evidence for involvement of RNA:DNA hybrids in DNA replication was presented in 1972 by Okazai and colleagues in their investigation of the discontinuous mechanism of DNA replication (Sugino et al., 1972). Subsequent studies showed that the RNA component of

the RNA:DNA hybrid is targeted for removal, by the RNA:DNA hybrid specific ribonuclease H (RNase H) (Itoh and Tomizawa, 1980; Masukata et al., 1987). Similarly, RNA:DNA hybrids have been shown to prime DNA synthesis from bacteriophage T4 replication origins following infection (Carles-Kinch and Kreuzer, 1997). In later years, the involvement of RNA:DNA hybrids has been extended to synthesis of leading strand, where the RNA component of the hybrid provides a primer for DNA polymerases in species ranging from *Saccharomyces cerevisiae* (Baldacci et al., 1984), to mammals (Pohjoismäki et al., 2010; Xu and Clayton, 1995). Interestingly, by overlapping DNA replication origins in human cells with published maps of R-loop distribution in mammals, recently, Gómez and colleagues proposed that human cells could also employ RNA:DNA hybrids together with G-quadruplexes for the initiation of DNA replication (Lombraña et al., 2015), although clear experimental support for R-loop role in initiation of DNA replication remain to be seen (Ekundayo and Bleichert, 2019).

### 1.5.4. The role of RNA:DNA hybrids in *Ig* class switch recombination.

In response to stimulation (e.g. through a contact with a foreign antigen), B-cells undergo a process called immunoglobulin (Ig) class switch recombination (CSR) that is responsible for switching B-cell antibody production from IgM (immunoglobulin M) to IgG, IgE or IgA (Chaudhuri and Alt, 2004; Stavnezer et al., 2008). At the molecular level, CSR occurs within a G-rich non-template DNA strand of 25–80 bp tandem repeats stretching over several kilobases, termed the S region (Dunnick et al., 1993). In stimulated B-cells, it is thought that non-coding (sterile) transcripts (Manis et al., 2002; Shinkura et al., 2003) are generated in the S-region where they form stable RNA:DNA hybrids (Daniels and Lieber, 1995; Reaban and Griffin, 1990). In a seminal paper by Lieber and colleagues, the presence of RNA:DNA hybrid in the S-region of stimulated B-cell was not only confirmed to take place *in vivo* but also shown to be part of stable R-loops (Yu et al., 2003). As a result, the G-rich non-template, single stranded DNA, displaced by the formation of R-loops in the S-region, becomes a target for

deamination by the activation-induced cytidine deaminase (AID) (Chaudhuri et al., 2003). In turn, AID-mediated deamination of cytosine bases to deoxyuridine is thought to mediate double strand DNA break (Catalan et al., 2003; Faili et al., 2002; Petersen-Mahrt et al., 2002). The deoxyuridine is targeted for removal by uracil DNA glycosylase (UNG), a component of the base excision repair (BER) (Petersen-Mahrt et al., 2002). UNG processing of deoxyuridines has been shown to result in an abasic site whose processing by the apurinic/apyrimidinic endonuclease 1 (APE1) is found sufficient for nick formation that precede CSR (Chaudhuri and Alt, 2004; Imai et al., 2003; Kao et al., 2013; Petersen-Mahrt et al., 2002). Ironically, despite the essential role of R-loops in mediating CSR, failure to properly regulate the formation of R-loops or untimely removal has been shown to compromise the genomic integrity of B-cells itself (Chaudhuri and Alt, 2004; Pefanis et al., 2014). These threats to genomic stability posed by the presence of R-loops is predominantly driven by the vulnerability of the displaced single stranded DNA to genotoxic attack (Crossley et al., 2019; Lindahl, 1993; Santos-Pereira and Aguilera, 2015).

### 1.5.5. The role of RNA:DNA hybrids in transcriptional regulation

Although transcription is necessary to generate R-loops, accumulating evidence suggest that transcription itself is subject to regulation by R-loops (Aguilera and García-Muse, 2012; Crossley et al., 2019). First *in vivo* indication of R-loops' role in transcriptional regulation came from studies of topoisomerase 1 deficient *Escherichia coli* cells (Drolet, 2006; Drolet et al., 1995). In these bacterial cells, negative supercoiling increased as a result of topoisomerase deficiency was found to promote excessive R-loops formation and growth defect that can be rescued by overexpression of RNase H, R-loops specific ribonuclease (Drolet, 2006; Drolet et al., 1995). Studies on the impact of excessive R-loops formation leading to hyper-negative supercoiling in *Escherichia coli*, and *Saccharomyces cerevisiae* showed that, unless resolved, it compromises fidelity of transcription (Baaklini et al., 2008; El Hage et al., 2010).

Subsequently, work from the Proudfoot and Gromak labs showed that R-loops formation in the G-rich sequences of 3'end in mammalian genes, contribute to efficient transcriptional termination by pausing RNA polymerase II (Gromak et al., 2006; Skourti-Stathaki et al., 2011). In addition to promoting transcriptional termination, R-loops formed at 3'-end of highly transcribed convergent genes are thought to prevent transcriptional read through into the adjacent locus (Crossley et al., 2019; Sanz et al., 2016a). Consistent with this, recent evidence suggest that formation of R-loops also acts as a strong barrier for RNA polymerase II mediated transcription (Belotserkovskii et al., 2017), potentially contributing to transcriptional stress (Saponaro et al., 2014). While promoting transcription termination *in cis*, R-loops have also been suggested to promote antisense transcription from the displaced single stranded DNA (Nair et al., 2020; Skourti-Stathaki et al., 2014; Tan-Wong et al., 2019). This is thought to lead to the formation of double stranded RNA at transcriptional pause sites, ultimately leading to the formation of repressive chromatin marked by H3K9me2 (Skourti-Stathaki et al., 2014). However, genome-wide R-loop mapping by the Chedin lab did not seem to corroborate the presence of repressive histone marks in R-loops harbouring genomic regions (Sanz et al., 2016a). Similarly, the presence of RNA:DNA hybrids has also been shown to disfavour the presence of histone proteins (Dunn and Griffith, 1980; Crossley et al., 2019). In addition to promoting transcriptional termination (Skourti-Stathaki et al., 2011), recent genome-wide mapping of R-loops using a variety of methods demonstrate enrichment of R-loops at promoters of mammalian genes (Chen et al., 2017; Ginno et al., 2012; Nadel et al., 2015; Sanz et al., 2016a). Such promoter associated R-loops have been suggested to counter promoter DNA methylation (Ginno et al., 2012) by hindering binding of methyltransferases to unmethylated CpG islands (Grunseich et al., 2018; Ginno et al., 2012). Correspondingly, excessively accumulated R-loops were associated with hypomethylation in Aicardi-Goutières syndrome genome that lack optimal R-loop resolution mechanism (Lim et al., 2015). This

hypomethylated feature is thought to render R-loop harbouring promoters accessible for transcription activation (Boque-Sastre et al., 2015; Cloutier et al., 2016; Arab et al., 2019), potentially through recruitment of enzymes involved in active DNA demethylation (Arab et al., 2019). Paradoxically, R-loops have also been shown to impede transcription elongation and serve as sites of recombination (Huertas and Aguilera, 2003; Domínguez-Sánchez et al., 2011; Aguilera and García-Muse, 2012). These features of R-loop formation necessitate tight balance between their rate of deposition and removal to prevent potential deleterious consequences.

### 1.5.6. RNA:DNA hybrids and genome instability

Despite the important role R-loops play in physiological processes ranging from transcriptional regulation to class switch recombination in B-cells, persistent R-loops formation can also pose significant threat to genome instability. The earliest evidence for R-loops involvement in genome instability came from studies of *Saccharomyces cerevisiae* mutants of the THO complex (Huertas and Aguilera, 2003). THO is a multimeric protein complex composed of HPR1, MFT1 and THP2 proteins that mediate nuclear export of polyadenylated mRNAs in a transcription dependent manner (Chávez et al., 2000), in association with the TREX complex (Aguilera and Gómez-González, 2008). THO mutant *Saccharomyces cerevisiae* showed defective nuclear export, excessive R-loop accumulation and hyperrecombination (Huertas and Aguilera, 2003; Luna et al., 2005). Interestingly, the hyper-recombination was shown to be R-loop dependent as over-expression of the RNA:DNA hybrid ribonuclease H suppressed the hyper-recombination phenotype of THO mutants (Huertas and Aguilera, 2003; Jimeno et al., 2002; Luna et al., 2005). Subsequently, mutation of THO/TREX has been shown to lead to R-loop accumulation-mediated genome instability in humans (Domínguez-Sánchez et al., 2011), suggesting a conserved THO/TREX function complex in safeguarding genome stability by limiting excessive R-loops formation (Santos-Pereira and Aguilera, 2015). In addition to their role in transcription-associated recombination (Belotserkovskii et al., 2017; Crossley et al.,

2019), persistent R-loops have been shown to cause genome instability in variety of other processes (Santos-Pereira and Aguilera, 2015; Costantino and Koshland, 2018). These include, forming physical impediment to progression of replication fork in dividing cells (Gan et al., 2011; Kotsantis et al., 2016; Stork et al., 2016; Wellinger et al., 2006), especially in head-on collision with replisome (Hamperl et al., 2017; Lang et al., 2017; Matos et al., 2020), DNA fragmentation (Li and Manley, 2005), activation-induced deaminase (AID) mediated hypermutation from the non-template single stranded DNA (Petersen-Mahrt et al., 2002; Chaudhuri et al., 2003; Imai et al., 2003; Kao et al., 2013), as well as exacerbating DNA damage repair by impeding the binding of DNA damage repair proteins to sites of double stranded DNA breaks (Aguilera and Gómez-González, 2008; Cohen et al., 2018; Gorthi et al., 2018; Quan Zhu et al., 2018; Barroso et al., 2019). Paradoxically, R-loops have also been implicated in mediating DNA damage repair. For example, Fischer and colleagues showed that transient RNA:DNA hybrids are essential for homologous recombination-mediated DNA damage repair in *yeast* genome as depletion of transient hybrids via RNase H over-expression led to strand resection (Ohle et al., 2016). Similarly, R-loops mediated recruitment of Rad52 have been shown to mediate transcription-associated homologous recombination repair of double strand breaks in human cells (Yasuhara et al., 2018), suggesting a conserved mechanism of R-loops-mediated double strand break repair. Moreover, DROSHA-mediated RNA:DNA hybrid formation around DNA double strand break sites is crucial for error-free repair (Lu et al., 2018). The apparently contradictory roles of RNA:DNA hybrids in double strand breaks formation and repair suggest the existence of cell-type specific mechanisms to maintain a tight lid on physiological DNA repair-related RNA:DNA hybrids, which are likely to be short lived (Depamphilis et al., 2012), from their transitioning to persistent pathological R-loops (Crossley et al., 2019; García-Muse and Aguilera, 2019).

### 1.5.7. Mechanisms of RNA:DNA hybrid resolution

Given the threat persistent R-loops can pose to genomic stability (Crossley et al., 2019; García-Muse and Aguilera, 2019), eukaryotic genomes deploy a variety of surveillance mechanisms to either prevent the unscheduled formation of R-loops or ensure their rapid dissolution once they formed (Skourti-Stathaki and Proudfoot, 2014). Since defective transcription and transcript processing have been shown to contribute to R-loops formation, there is increasing realisation that one of the ways organisms keep a lid on R-loop levels is to prevent their formation in the first place (Kim and Jinks-Robertson, 2012; Wahba et al., 2011), for example, by ensuring the precise spatial and temporal regulation of transcriptional events that can contribute to R-loop formation (Paulsen et al., 2009; Santos-Pereira and Aguilera, 2015), starting from initiation (Dumelie and Jaffrey, 2017) to elongation (Domínguez-Sánchez et al., 2011; Zatreanu et al., 2019), splicing (Li and Manley, 2005; Nguyen et al., 2019), nuclear export and degradation (Aguilera and García-Muse, 2012; Huertas and Aguilera, 2003). Despite this, nascent RNA has been shown to exhibit strong tendency to re-anneal to excessively negative supercoiled DNA template (Drolet et al., 1994), as a result, organisms deploy specific proteins and modification of local chromatin structure to keep in check levels and duration of transcription associated hyper-negative supercoiling to limit R-loop formation (Tuduri et al., 2009; El Hage et al., 2010; Cristini et al., 2019; Fragola et al., 2020); (Bayona-Feliu et al., 2017; Zeller et al., 2016). In case any of the preventative mechanisms fail, and unscheduled R-loops build-up as a result, organisms from bacteria to human developed a specific class of ribonucleases called RNase H (Cerritelli and Crouch, 2009; Hyjek et al., 2019). In eukaryotic organisms, there are two types of RNase H enzymes, RNase H1 and RNase H2 and both of them catalyse dissolution of RNA:DNA hybrids by endonucleolytical degradation of the RNA component of the hybrids (Cerritelli and Crouch, 2009; Skourti-Stathaki and Proudfoot, 2014). Whereas RNase H1 acquired the ability to both

bind the RNA:DNA hybrid and degrade the RNA part of the hybrid (Gaidamakov et al., 2005), RNase H2 is made-up of three different protein subunits, RNase H2A that comprise the catalytic subunit, RNase H2B, and RNase H2C (Reijns et al., 2012; Zimmer and Koshland, 2016; Hyjek et al., 2019). Unlike RNase H2, RNase H1 has been shown to bind to the majority of nuclear R-loops *in vivo* (Zimmer and Koshland, 2016; Chen et al., 2017) suggesting differences in their R-loop targeting efficiencies (Lockhart et al., 2019). In addition to RNase H proteins, eukaryotic genomes use the SETX, a putative senataxin encoded RNA:DNA helicase to resolve RNA:DNA hybrids (Kim et al., 1999) in species ranging from *yeast* (Mischo et al., 2011) to humans (Skourti-Stathaki et al., 2011). More recently, senataxin has been implicated in the resolution of R-loops formed at sites of double strand DNA breaks (Cohen et al., 2018; Bader and Bushell, 2020), suggesting a possible function for this enzyme in the repair of DNA double strand breaks (Skourti-Stathaki and Proudfoot, 2014). In addition, a family member of PIF1 DNA helicase, Pif1p has also been implicated in the resolution of RNA:DNA hybrid structures (Boulé and Zakian, 2007). Defining R-loops resolution by the ability of a given protein to suppress γH2AX (a central mark of DNA damage) signal in eukaryotes (Georgoulis et al., 2017), recent genome-wide screens (Paulsen et al., 2009) and proteomic analysis (Cristini et al., 2018; Wang et al., 2018) indicate the presence of hundreds of previously unknown factors potentially involved in R-loop resolution and genome stability (Hegazy et al., 2020), yet many of these remain completely unexplored.

### 1.5.8. Detection and mapping of RNA:DNA hybrids

Recognising the diverse biological functions that R-loops play in transcriptional regulation (Santos-Pereira and Aguilera, 2015) and genome stability (Aguilera and Gómez-González, 2008; Paulsen et al., 2009; Costantino and Koshland, 2018; García-Muse and Aguilera, 2019), several strategies that aim to detect RNA:DNA hybrids have been developed (Boque-

Sastre et al., 2017; Halász et al., 2017; Vanoosthuyse, 2018). The most widely used RNA:DNA hybrid mapping strategy involves the use of S9.6, a monoclonal anti-RNA:DNA hybrid antibody (Boguslawski et al., 1986) that has high specificity for RNA:DNA hybrids over other nucleic acid structures (Boguslawski et al., 1986; Phillips et al., 2013). Using sensitivity of R-loops structures to RNase H treatment as a control, S9.6 antibody has been used to study R-loops as part of several techniques, including dot-blot (Wahba et al., 2013), immunofluorescence staining (García-Rubio et al., 2015; Stork et al., 2016; Hamperl et al., 2017; Barroso et al., 2019), and affinity purification of R-loop structures (Chan et al., 2014; El Hage et al., 2014; Nadel et al., 2015; Cristini et al., 2018) often followed by high-throughput sequencing for genome-wide mapping of R-loops (Ginno et al., 2012; Sanz et al., 2016a; Wahba et al., 2016; Cristini et al., 2018; Kabeche et al., 2018). In addition to detection of R-loops using S9.6 antibody, recently, a catalytically inactive RNase H has been used for mapping genome-wide profiles of R-loops (Ginno et al., 2012; Chen et al., 2017). Moreover, compared to double stranded DNA, the RNA comprising R-loops increases molecular weight of sequences harbouring R-loops, and as a result, R-loop-containing sequences has been shown to have a distinct electrophoresis in electrophoretic mobility shift (EMSA) analysis (Daniels and Lieber, 1995). Accordingly, EMSA has been used to study R-loops in the S region of stimulated B-cells during class switch recombination (Yu et al., 2003). Alternatively, R-loops can be mapped at single nucleotide resolution by exposing the displaced single stranded DNA across R-loop containing regions to sodium bisulfite treatment, which converts cytidine bases to uracils, which are identified as thymine bases by PCR (Yu et al., 2003, 2006; Malig et al., 2020). Taking clues from sequences known to be prone to R-loop formation, recently, a computational model for prediction of R-loop detection has been proposed (Jenjaroenpun et al., 2015). Despite the various ways R-loops

can be mapped, there is a growing appreciation of limitations of these techniques (Halász et al., 2017; Hartono et al., 2018; Vanoosthuyse, 2018).

### 1.6. Aims of the thesis

Despite the knowledge of N6-methyladenosine (m6A) deposition on RNAs for decades, interest in the potential biological functions of m6A, from development to disease, has only intensified following the recent discovery of enzymes demonstrated its reversible and dynamic nature. As a result, m6A has been implicated in almost all steps of RNA metabolism. Meanwhile, the DNA counterpart of m6A, N6-methyldeoxyadenosine (6mA) that was thought to be confined to prokaryotic and unicellular eukaryotic genomes, recently, has also been discovered in *Drosophila*, *Caenorhabditis elegans*, and *Chlamydomonas reinhardtii* genomes. Although these findings led to multiple searches for 6mA in higher eukaryotes, due to technical limitations often combined with insufficiently controlled experiments, the results of these initial attempts proved contradictory, thus further propagating uncertainty regarding 6mA presence in vertebrate genomes.

#### 1.6.1. Investigate the potential presence of N6-methyladeoxyadenosine (6mA) in the human genome.

The first and primary aim of this thesis was therefore to investigate the presence of 6mA in the genome of human pluripotent stem cells (hPSCs). As described in chapter 3.1 – 3.3, 6mA was not detectable neither using the highly sensitive immunostaining technique nor with the ultra-sensitive mass spectrometry.

#### 1.6.2. Investigate the origin of chromatin-associated 6mA/m6A immunostaining signal in human cells.

Although 6mA was undetected in human cells, paradoxically, the immunostaining analysis showed persistent 6mA/m6A signal refractory to treatments with RNase A or DNase I. As

described in chapter 3.4, using multiple independent approaches, this thesis uncovers the presence of RNase H-sensitive m6A-marked RNA:DNA hybrids in hPSCs.

### 1.6.3. Investigate the genome-wide distribution of m6A marked RNA:DNA hybrids in human cells

As m6A presence on RNA:DNA hybrids was not known before, there was no protocol for genome-wide mapping of m6A marked RNA:DNA hybrid profiles. Therefore, in chapter 4 of this thesis, a novel technique for genome-wide mapping of m6A-marked RNA:DNA hybrids is presented. Remarkably, at the sequence levels, there was an almost complete over-lap of m6A-decorated R-loops with S9.6 captured R-loops, suggesting that m6A marks most RNA:DNA hybrids in hPSCs.

### 1.6.4. Interrogate the relationship between m6A marked RNA:DNA hybrids and transcription.

The discovery that m6A marks the majority of RNA:DNA hybrids prompted the interrogation of transcriptional profiles of genomic sequences harbouring m6A marked RNA:DNA hybrids. In chapter 5.1. – 5.2

### 1.6.5. Investigate the relationship between histone modifications and m6A marked RNA:DNA hybrids in human pluripotent stem cells.

Previous investigations of S9.6 mapped R-loops suggested potential modulation of histone modifications in R-loop containing loci. As described in chapter 5.3, the relationship between m6A marked RNA:DNA hybrids and histone modifications is investigated. Interestingly, m6A marked RNA:DNA hybrid did not exhibit significant overlap with any of the known histone marks.

### 1.6.6. Investigate the cell-cycle dynamics of m6A marked RNA:DNA hybrids in human pluripotent stem cells.

Following the discovery of m6A on RNA:DNA hybrids, there was a persistently high m6A immunostaining signal in mitotic cells compared to interphase cells. As described in chapter 6, this prompted the investigation of m6A marked RNA:DNA hybrid dynamics in flow cytometry-sorted G0/G1, S, and G2/M cell-cycle phases. Interestingly, m6A-marked RNA:DNA hybrids showed preferential enrichment in S and G2/M cell-cycle phases compared to cells in G0/G1. Importantly, both m6A-marked RNA:DNA hybrids and S9.6 pulled-down R-loops exhibited identical cell-cycle dynamics.

### 1.6.7. Investigate the biological function of m6A on RNA:DNA hybrids.

One of the best characterized function of m6A on RNA is its ability to shorten messenger RNA lifespan. As described in chapter 7, this prompted modulation of m6A levels through targeted depletion of its writer enzyme, which led to genome-wide accumulation of RNA:DNA hybrids, suggesting m6A involvement in regulating the stability of RNA:DNA hybrids in human cells. In chapters 8 and 9, YTHDF2, one of the best characterised readers of m6A, is shown to co-localise with RNA:DNA hybrids *in vivo* and physically interact with m6A marked RNA:DNA hybrids *in vitro*. Importantly, both the transient and CRISPR-cas9-mediated depletion of YTHDF2 led to pronounced accumulation of RNase H-sensitive RNA:DNA hybrids, cell growth retardation and increased DNA damage in human and mouse cells, suggesting a conserved role for m6A in regulating the stability of RNA:DNA hybrids. The implications of these findings are discussed in chapter 10.

**1.7. List of publications:**

**Publications during PhD**

1- **Abdulkadir Abakir**, Tom C. Giles, Agnese Cristini, Jeremy M. Foster, Nan Dai, Marta Starczak, Alejandro Rubio-Roldan, Miaomiao Li, Maria Eleftheriou, James Crutchley, Luke Flatt, Lorraine Young, Daniel J. Gaffney, Chris Denning, Bjørn Dalhus, Richard D. Emes, Daniel Gackowski, Ivan R. Corrêa Jr, Jose L. Garcia-Perez, Arne Klungland, Natalia Gromak, Alexey Ruzov. (**2020**). N6-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. ***Nat Genet***, 52, pages48–55

- Featured in a 'news and views' article in Nature Genetics. Marnef, A., Legube, G. (2019). $m^6A$ RNA modification as a new player in R-loop regulation. *Nat Genet* https://doi.org/10.1038/s41588-019-0563-z.

- Highlights in media: https://phys.org/news/2020-01-rna-effect-dna.html

2- **Abdulkadir Abakir**, Alexey Ruzov. **(2020).** Analysis of 5-carboxylcytosine distribution using DNA Immunoprecipitation (DIP). ***Methods in Molecular Biology***. *In print*

3- **Abdulkadir Abakir**, Alexey Ruzov. **(2020)**. Detection of Low-Abundance DNA Modifications Using Signal Amplification-Based Immunocytochemistry. ***Methods in Molecular Biology***. *In print*

4- Seema Rajani*, Christopher Gell*, **Abdulkadir Abakir**, Robert Markus. **(2020)**. Computational Analysis of DNA Modifications in Confocal Images. ***Methods in Molecular Biology***. *In print*

5- Igor Yakovlev, Daniel Gackowski, **Abdulkadir Abakir**, Marcos Viejo, Alexey Ruzov, Ryszard Olinski, Marta Starczak, Carl Gunnar Fossdal, Konstantin V. Krutovsky. (**2019**). Mass spectrometry reveals the presence of specific set of epigenetic DNA modifications in the Norway spruce genome. ***Scientific Reports*** 9**,** 19314

6- Gackowski D, Gawronski M, Kerr C, Radivoyevitch T, Zarakowska E, Starczak M, **Abakir A**, Ruzov A, Maciejewski JP, Olinski R. (**2019**). 5-formylcytosine and 5-hydroxymethyluracil as surrogate markers of TET2 and SF3B1 mutations in myelodysplastic syndrome, respectively. ***Haematologica***

7- Ramsawhook AH, Lewis LC, Eleftheriou M, **Abakir A**, Durczak P, Markus R, Rajani S, Hannan NRF, Coyle B, Ruzov A (**2017**). Immunostaining for DNA Modifications: Computational Analysis of Confocal Images. ***J. Vis. Exp***. (127).

**Book Chapter:**

8- Moler ERV, **Abakir A**, Eleftheriou M, Johnson JS, Krutovsky KV, Lewis LC, Ruzov A, Whipple AV, Rajora OP (**2018**). Population Epigenomics: Advancing Understanding of

Phenotypic Plasticity, Acclimation, Adaptation and Diseases. In: Rajora O. (eds) Population Genomics. Population Genomics. ***Springer***

**Pre – PhD publications**

9- **Abdulkadir Abakir**, Lee Wheldon, Andrew Johnson, Patrick Laurent, Alexey Ruzov. (**2016**). Detection of Modified Forms of Cytosine Using Sensitive Immunohistochemistry**. *J. Vis. Exp***. (114).

10- Eleftheriou M, Pascual AJ, Wheldon LM, Perry C, **Abakir A**, Arora A, Johnson AD, Auer DT, Ellis IO, Madhusudan S, Ruzov A. (**2015**). 5-Carboxylcytosine levels are elevated in human breast cancers and gliomas. ***Clinical Epigenetics*** 7(88).

11- Wheldon LM*, **Abakir A*** Ferjentsik Z*, Dudnakova T, Strohbuecker S, Christie D, Dai N, Guan S, Foster JM, Corrêa IR Jr, Loose M, Dixon JE, Sottile V, Johnson AD, Ruzov A. (**2014**). Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. ***Cell Reports*** 7, 1353–1361. **Joint first authors*.

12- Alioui A, Wheldon LM, **Abakir A**, Ferjentsik Z, Johnson AD, Ruzov A. (**2012**). 5-Carboxylcytosine is localized to euchromatic regions in the nuclei of follicular cells in axolotl ovary. ***Nucleus*** 3(6), 565-9

**Book Chapter:**

**Abakir A**, Wheldon L.M, Ruzov A. (**2016**). Immunohistochemical Detection of Oxidized Forms of 5-Methylcytosine in Embryonic and Adult Brain Tissue. In: Karpova N. (eds) Epigenetic Methods in Neuroscience Research. Neuromethods, vol 105. ***Humana Press***, New York, NY.

.

# Chapter 2

# Materials and methods

## Chapter 2: Materials and methods

### 2.1. Materials:

#### 2.1.1. Cell lines and mouse tissues

REBL-PAT human induced pluripotent stem cells (hiPSCs) were generated previously from patient skin derived fibroblasts using CytoTune 2.0 nonintegrating Sendai virus (ThermoFisher) (Duncan et al., 2017). HUES7 hESCs were gifted by Chad Cowan and Doug Melton at the Harvard Stem Cell Institute (Cowan et al., 2004). HeLa, LN-18 and U87MG cancer cell lines were gifts from Dr Beth Coyle of Nottingham University, sourced from ATCC. YTHDF2 knockout (CRISPR/Cas9-mediated deletion of 140bp in exon3, resulting in frameshift and premature stop codon) HAP1 cells (Horizon Discovery, no. HZGHC006678c001) and their isogenic WT parental HAP1 cells were obtained from Horizon Descovery. Wild type and Ythdf2 KO mouse neural stem cells (mNSCs) together with wild type and Ythdf2 KO mouse embryonic cortex were previously derived from WT and Ythdf2 KO mouse E14.5 embryonic forebrain (Li *et al*., 2018).

#### 2.1.2. Antibodies

**Table 2.1.2**. List of primary and secondary antibodies used for immunofluorescence (IF), Western blot (WB), ChIP and DRIP procedures. The table shows the name of antibodies, supplier, catalogue number, and procedure(s) the antibody were used in throughout this thesis.

| Antibody | Supplier | Catalogue number | Procedure |
|---|---|---|---|
| Anti-m6A rabbit polyclonal | Synaptic systems | 202003 | IF/DRIP/WB |
| Anti-DNA-RNA Hybrid S9.6 antibody | Merck Millipore | MABE1095 | IF/DRIP/WB |
| Anti-m1A mouse monoclonal | Diagenode | C15200235 | IF |

| | | | |
|---|---|---|---|
| Anti-5mC mouse monoclonal | Diagenode | C15200081-100 | IF |
| Anti-phospho-Histone H3 (Ser10) mouse monoclonal | Cell Signalling | (6G3) 9706 | IF |
| Anti-5hmC mouse monoclonal | Diagenode | C15220001-50 | IF |
| Anti-YTHDF1 rabbit polyclonal | ProteinTech | 17479-1-AP | IF/WB |
| Anti-YTHDF2 rabbit polyclonal | ProteinTech | 24744-1-AP | IF/ChIP/WB |
| Anti-HNRNPA2B1 mouse monoclonal | Novus | NB120-6102SS | IF/ChIP |
| Anti-γH2AX mouse monoclonal | Merck | 05-636 | IF/ChIP |
| peroxidase-conjugated anti-rabbit secondary antibody | ThermoFisher | 65-6120 | IF |
| Alexafluor 633-conjugated secondary antibody | ThermoFisher | A-21052 | IF |
| Top1 | Abcam | ab109374 | WB |
| METTL3 rabbit polyclonal | Bethyl Laboratories | A301-567A | WB |
| Lamin B1 rabbit polyclonal | Abcam | ab16048 | WB |

### 2.1.3. Immunofluorescence reagents

**Table 2.1.3. Immunofluorescence reagents for staining mammalian cells and paraffin embedded tissues. Details of antibodies used are provided in Table 2.1.2.**

| Reagent/material | Supplier | Catalogue number |
|---|---|---|
| 8-well glass chamber slides | ThermoFisher | 154534PK |

| | | |
|---|---|---|
| Glass Coplin jar | Cole-Parmer | UY-48585-30 |
| Xylene | ThermoFisher | 6601 |
| Ethanol, anhydrous denatured, histological grade | Sigma Aldrich | 64-17-5 |
| Phosphate buffer saline (PBS) | Fisher Scientific | 12821680 |
| 4% paraformaldehyde (PFA) | Sigma Aldrich | P6148 |
| Tween 20 | Sigma Aldrich | P9416 |
| Triton X-100 | Sigma Aldrich | X100 |
| Slide moisture chamber | Scientific Device Laboratory | 197-BL |
| BSA (blocking solution) | Sigma Aldrich | A9418 |
| 22 x 32 mm Glass cover slips | BDH | 406/0188/24 |
| Mounting Medium with DAPI | Vector Labs | H-1200 |
| Nail polish | Cosmetic retailers (e.g. Boots) | Colourless |
| Hydrochloric Acid, 37% | Fisher Scientific | 7647-01-0 |
| RNase A | Qiagen | 19101 |
| RNase H | NEB | M0297S |
| Turbo DNase | Invitrogen | AM1907 |
| Tyramide signal amplification Kit Plus Fluorescein System | Perkin Elmer | NEL741001KT |

**2.1.4. Cell culture reagents**:

**Table 2.1.4.1. Human pluripotent stem cell culture reagents. This includes reagents for culturing both REBL-PAT hiPSCs and HUES7 hESCs.**

| Reagent | Supplier | Catalogue number |
|---|---|---|
| Essential 8™ medium | LifeTechnologies | A1517001 |
| Corning® Matrigel® | Corning | 356235 |
| Phosphate-buffered saline (PBS) | Gibco™ | 14190-094 |
| TrypLE™ Select Enzyme | ThermoFisher | 12563029 |
| ROCK inhibitor | Tocris Bioscience | 1254/10 |
| Non-Essential Amino Acids | Gibco™ | 11140050 |
| 100 μM β-mercaptoethanol | Sigma | 63689 |
| GlutaMAX | Gibco™ | 35050061 |
| DMEM-F12 | Gibco™ | 11320033 |
| Knockout Serum Replacement | Gibco™ | 10828028 |
| KaryoMAX™ Colcemid™ Solution | ThermoFisher | 15212012 |
| Basic Fibroblast Growth Factor | Peprotech | 100-18B |

**Table 2.1.4.2. Cell culture reagents for HeLa, LN-18 and U87MG cancer cell lines.**

| Reagent | Supplier | Catalogue number |
|---|---|---|
| DMEM | Gibco™ | 11965092 |
| Fetal Bovine Serum (FBS) heat inactivated | ThermoFisher | F4135 |
| Penicillin-Streptomycin (P/S) | Gibco™ | 15140122 |
| Phosphate-buffered saline (PBS) | Gibco™ | 14190-094 |
| Trypsin-EDTA (0.25%) | Gibco™ | 25200056 |

**Table 2.1.4.3. Cell culture reagents for wild type and YTHDF2 KO HAP1 cells.**

| Reagent | Supplier | Catalogue number |
|---|---|---|
| DMEM/F12 | Gibco™ | 11320033 |
| Fetal Bovine Serum (FBS) heat inactivated | ThermoFisher | F4135 |
| Penicillin-Streptomycin (P/S) | Gibco™ | 15140122 |
| Phosphate-buffered saline (PBS) | Gibco™ | 14190-094 |
| Trypsin-EDTA (0.25%) | Gibco™ | 25200056 |
| Attachment Factor Protein | Gibco™ | S006100 |

### 2.1.5. Flow cytometry materials and reagents to sort human cells into G0/G1, S and G2/M phases of the cell cycle.

| Material and reagents | Supplier | Catalogue number |
|---|---|---|
| RNase A (100µgml) | Qiagen | 19101 |
| Phosphate-buffered saline (PBS) | Gibco™ | 14190-094 |
| Ethanol, anhydrous denatured, histological grade | Sigma Aldrich | 64-17-5 |
| Propidium iodide (10µgml) | Sigma-Aldrich | P3566 |
| MoFlo Astrios EQ | Beckman | B25982 |
| Beckman Coulter Kaluza | Beckman | 2.1 sofware |

### 2.1.6. Chromatin immunoprecipitation reagents

| Materials and reagents | Supplier | Catalogue number |
|---|---|---|
| Formaldehyde (36.5-38%) | Sigma-Aldrich | F8775 |
| EZ-Magna ChIP A/G ChIP kit | Merck | 17-10086 |
| Covaris S2 ultrasonicator | Covaris | S220 |

| | | |
|---|---|---|
| Anti-YTHDF2 rabbit polyclonal antibody | ProteinTech | 24744-1- AP |
| anti-HNRNPA2B1 mouse monoclonal antibody | Novus | NB120-6102SS |
| anti-γH2AX mouse monoclonal antibody | Merck | 05-636, clone JBW301 |

2.1.7.1. **Materials and reagents for m6A and S9.6 DRIP as well as m6A RIP**

| **Materials and reagents** | **Supplier** | **Catalogue number** |
|---|---|---|
| DNeasy Blood & Tissue Kit | Qiagen | 69504 |
| Kapa Library Quantification Kit | Kapa Biosystems | KK4823 |
| The Agilent 2200 TapeStation system | Aiglent Technologies | |
| Phenol:Chloroform:Isoamyl Alcohol 25:24:1 | Aldrich | P2069-100ML |
| RNeasy mini Kit | Qiagen | 74104 |
| NanoDrop™ 2000/2000c | ThermoFisher | ND-2000 |
| Proteinase K (PK) Solution | Promega | MC5005 |
| Phosphate-buffered saline (PBS) | Gibco™ | 14190-094 |
| DynaMag™-2 Magnet | ThermoFisher | 12321D |
| SuperScriptIII Reverse Transcriptase | Invitrogen | 2072529A |
| Dynabeads™ M-280 Sheep Anti-Rabbit IgG | Invitrogen | 11203D |
| Triton X-100 | Sigma Aldrich | X100 |
| EpiMark N6-Methyladenosine Enrichment Kit | NEB | E1610S |
| Anti-m6A rabbit polyclonal antibody | Synaptic systems | 202003 |

| | | |
|---|---|---|
| Random hexamers | Invitrogen | 1831815 |
| Dynabeads™ M-280 Sheep Anti-Mouse IgG | Invitrogen | 11201D |
| Anti-DNA-RNA Hybrid S9.6 antibody | Merck Millipore | MABE1095 |
| RNase H | NEB | M0297S |
| RNase A | Qiagen | 19101 |
| Turbo DNase | Invitrogen | AM1907 |
| Covaris S2 ultrasonicator | Covaris | S220 |
| QIAquick PCR purification kit | Qiagen | 28104 |
| Nuclease free water | NEB | B1500S |
| Qubit™ Flex Fluorometer | Invitrogen | Q33327 |
| NEBNext® Multiplex Oligonucleotides for Illumina ® | NEB | E7335S |
| NEB Next DNA Ultra Library Preparation Kit | NEB | E7370 |

### 2.1.7.2. **Buffers for m6A DIP and S9.6 DRIP:**

| **Buffer** | **Composition** |
|---|---|
| Lysis buffer | 20 mM Tris pH8.0, 4mM EDTA, 20mM NaCl, 1% SDS |
| 10x IP buffer | 100 mM Na-phosphate pH 7.0 (monodibasic), 1.4 M NaCl, 0.5 % Triton X-100. |
| TE buffer | 10mM Tris pH7.5, 1mM EDTA. |
| PBS (1X) | 137 mM NaCl, 2.7 mM KCl, 8 mM Na2HPO4, and 2 mM KH2PO4; pH 7.5. |

| | |
|---|---|
| Proteinase K buffer | 20 mM Tris pH 8.0, 4 mM EDTA, 20 mM NaCl, 1 % SDS. |

### 2.1.8. Oligonucleotides:

#### 2.1.8.1. Primers used for RT-qPCR (For assessing the expression levels of the corresponding loci)

| Oligonucleotide name | Sequence |
|---|---|
| *METTL3* forward | CAAGCTGCACTTCAGACGAA |
| *METTL3* reverse | GCTTGGCGTGTGGTCTTT |
| *YTHDF2* forward | ACCTTACTTGAGTCCACAGG |
| *YTHDF2* reverse | AGCCAATGGAGGGACTGTAG |
| *HNRNPA2B1* forward | AGGTGCTCCTCGCAGAGT |
| *HNRNPA2B1* reverse | TCCTAGAACTCTGAACTTCCTGC |
| *RNAse H1* forward | CAGGAATCGGCGTTTACTGG |
| *RNAse H1* reverse | TCACCTCTTTCCCTGCACTT |
| *GAPDH* forward | GATGCTGGCGCTGAGTACG |
| *GAPDH* reverse | GCAGAGATGATGACCCTTTTGG |
| *HECW1* forward | ATACAGCTACAACCCCGACC |
| *HECW1* reverse | CTGAGAGTGCCCGATGGAAT |
| *MAN2B2* forward | GTCCCACATCCCTTTCTCCA |
| *MAN2B2* reverse | GAACTGCTTGTCACATCCCC |
| *RANBP17* forward | CAACCTTCTACACAGCGCTC |
| *RANBP17* reverse | GCCCGATCAACATACGCTTT |

**2.1.8.2. Primers used for verifying the genotype of Wild type and CRISPR cas9 mediated YTHDF2 KO HAP1 cells.**

| Primer | sequence |
|---|---|
| *YTHDF2* 3rd exon forward | AATTATGCTTATGCACCTAGCTCCT |
| *YTHDF2* 3rd exon reverse | ATGCCATTCTTGGTCTTCAGTTTAG |

**2.1.8.3. Primers used for investigating enrichment in m6A DIP, S9.6 DRIP and ChIP experiments.**

| Primer | sequence |
|---|---|
| LINE1 forward | AGGCCACTGTGTGCGCGC |
| LINE1 reverse | CCAGGTGTGGGATATAGTCTC |
| Alu-Y forward | AGATCGAGACCATCCTGGCT |
| Alu-Y reverse | CCGCCTCCCGGGTTCACGCC |
| Alu-S forward | GCCGAGGCGGGCGGATCACC |
| Alu-S reverse | GCCTCCCGAGTAGCTGGGAT |
| α-satellites | Novus, NBP1-71654SS |
| MYH16 forward | AGAGCAGAGCAACAGAGACT |
| MYH16 reverse | TTCCCTACTGAGTCCTCCCA |
| HECW1 forward | CCCTTGACTCCCTTTCTCCC |
| HECW1 reverse | AGGGAGGGGAAGAAAGAGAG |
| MAD1L1 forward | TCACCACCATTCTCACGTCA |
| MAD1L1 reverse | TGATGGTGTGGAGAGTGGTC |

| | |
|---|---|
| MAN2B2 forward | CCACTGCTATGACCACCACT |
| MAN2B2 reverse | ACAGTGATGGGTGGTGATCA |
| LPA forward | AGTATTTTGTGGTTTTGATGCCA |
| LPA reverse | AACCCAGCAAGCTTAAACGC |
| RANBP17 forward | CACACGCAGAACCACACATA |
| RANBP17 reverse | CAGAAGCTAAGATGTGTGTGTGT |
| RANBP17 Downstream F | GGTGGTGGAAATTAGCAGGC |
| RANBP17 Downstream R | GGCAGAGGAGTTTAGTGACCA |
| HECW1 Downstream F | TGTCCTACCCCTACCCAAGA |
| HECW1 Downstream R | TTGGAGGGCTAAAGGACAGG |

**2.1.8.4. Oligonucleotides used for spike-in during m6A and S9.6 DRIP experiments:**

**These oligonucleotides were used to make synthetic RNA:DNA hybrids in which the RNA component either m6A modified or non-modified. Importantly, the DNA oligo was designed to include flanking primer sequences for assessment by qPCR.**

| Name | Sequence |
|---|---|
| m6A-RNA oligo | CCUAUGAGCAACGUGUUAGAUGGGCCGUUCAUCUGCUAAAA GG**X**CUGCUUUUGGGGCUUGUGCAAAGAUCAAACAACUUGU, where X is m6A |
| non-modified RNA oligo | CCUAUGAGCAACGUGUUAGAUGGGCCGUUCAUCUGCUAAAA GGACUGCUUUUGGGGCUUGUGCAAAGAUCAAACAACUUGU |

| | |
|---|---|
| DNA oligo | GGATACTCGTTGCACAATCTACCCGGCAAGTAGACGATTTT CCTGACGAAAACCCCGAACACGTTTCTAGTTTGTTGAACA |
| Amplification primer F | ACAAGTTGTTTGATCTTTGC |
| Amplification primer R | CCTATGAGCAACGTGTTAG |

**2.1.8.4. Oligonucleotides used in electrophoretic mobility-shift assays (EMSA). These were to assess affinity of YTHDF2 to m6A modified and non-modified RNA:DNA hybrids compared to its affinity to single stranded m6A modified and nonmodified RNA oligonucleotides.**

| **Name** | **Sequence** |
|---|---|
| Non-modified RNA oligo | AUGGGCCGUUCAUCUGCUAAAAGGACUGCUUUUGGGGCUUGU |
| m6A-containing RNA oligo | AUGGGCCGUUCAUCUGCUAAAAGG**X**CUGCUUUUGGGGCUUGU where X is m6A |
| DNA oligo | ACAAGCCCCAAAAGCAGTCCTTTTAGCAGATGAACGGCCCAT |

**2.1.8.5. Buffers used in EMSA**

| **Buffer** | **Composition** |
|---|---|
| High salt binding buffer | 12.5 mM Hepes pH 7.9, 5 mM MgCl2, 0.5 mM EDTA, 80/150 mM NaCl, 1 mM DTT, 0.025 % NP-40, 50 µg/ml BSA, 15 % glycerol |

| | |
|---|---|
| Low salt binding buffer | 6 mM Hepes pH 7.9, 4 mM Tris HCl pH 8.0, 30 mM KCl, 1 mM EDTA, 1 mM DTT, 30 µg/ml BSA, 5% glycerol |

### 2.1.9. siRNAs, transfection reagents, bacteria, plasmids

| Material and reagents | Supplier | Catalogue number |
|---|---|---|
| Human *METTL3* siRNA | Dharmacon | 56339 |
| Human *HNRNPA2B1* siRNA | ThermoFisher | 4390824 |
| Human *YTHDF2* siRNA | Qiagen | GS51441 |
| Human *RNase H1* siRNA | Dharmacon | M-012595-00-0010 |
| Non-targeting siRNA | ThermoFisher | D-001210-02 |
| DharmaFECT | Dharmacon | T-2001-01 |
| pEGFP-RNASEH1 plasmid | Addgene | 108699 |
| pmaxGFP | Lonza | VDC-1040 |
| FuGENE® HD | Promega | E2312 |
| DH5α Competent Cells | Invitrogen | 18265017 |
| QIAprep Spin Miniprep Kit | Qiagen | 27104 |

### 2.1.10. PCR, RT-PCR, and RT-qPCR

**Table 2.1.10. Reagent was used for PCR, RT-PCR, and RT-qPCR experiments.**

| Material and reagents | Supplier | Catalogue number |
|---|---|---|
| SuperScriptIII Reverse Transcriptase | Invitrogen | 2072529A |
| Nuclease free water | NEB | B1500S |
| RNeasy mini Kit | Qiagen | 74104 |

| | | |
|---|---|---|
| RNase-Free DNase Set | Qiagen | 79254 |
| DTT (0.1M) | ThermoFisher | P2325 |
| 1st strand synthesis kit | ThermoFisher | 18080400 |
| Phusion® High-Fidelity DNA Polymerase | NEB | M0530S |
| Random hexamers | Invitrogen | 1831815 |
| dNTPs | NEB | N0447S |
| SYBR Green PCR Master Mix | Sigma-Aldrich | S4438 |
| ABI 7500 Real-Time PCR system | Applied Biosystems | 4351105 |

### 2.1.11. Microscopes

| |
|---|
| Zeiss LSM780 Confocal microscope |
| Zeiss LSM 710 Confocal microscope (DIC) |
| Nikon Eclipse 90i wide-field microscope |

## 2.2. Cell culture experiments

**Cell culture was carried out in a mycoplasma free Class II safety cabinet.**

### 2.2.1. Human pluripotent stem cells

REBL-PAT human induced pluripotent stem cells (hiPSCs) and HUES7 human embryonic stem cells (hESCs) were cultured in Essential 8 medium with supplement (no. A1517001) on Matrigel-coated tissue culture fasks at 37 °C with 5% CO2. Approximately, every 3 days the cells were passaged by replacing E8 medium with PBS to wash the cells, followed by incubation with TrypLE Select Enzyme at 37°C in a humidified incubator with 5% CO2 to dissociate the cells. After 4 minutes, TryPLE was inactivated with BGK solution (DMEM-F12

medium supplemented with 100 μM β-mercaptoethanol, 1% GlutaMAX, 1% Non-Essential Amino Acids, 1% GlutaMAX and 15% Knockout Serum Replacement). Cells were then pelleted by centrifugation for 4 min at 160g, resuspendend in freshly prepared E8 medium supplemented with 10μM Rocki before seeding them into a freshly prepared Matrigel-coated tissue culture flask at approximately 25000 cells/$cm^2$. The E8 medium is replaced every 24hrs without Rocki supplementation until next split. Culture is restarted with a frozen batch of low passage cells after 20-30 passages. Use of the HUES7 hESC line was approved by the UK Medical Research Council Steering Committee in association with the UK Stem Cell Bank.

### 2.2.2. KaryoMAX Colcemid treatment

To obtain mitotic cell-enriched culture, hPSCs were treated with a 1:100 dilution of KaryoMAX Colcemid Solution when the cells were at their peak doubling time (24-48 hrs following seeding) and incubated at 37 °C in a humidified incubator with 5% CO2. After 3 hrs, the KaryoMAX containing medium was discarded and the cells were replenished with fresh KaryoMAX free E8 medium and incubated for 24hrs at 37 °C in a humidified incubator with 5% CO2. The cells were collected for either staining or DRIP experiments.

### 2.2.3. Fluorescence-activated cell sorting (FACS)

Approximately 70% confluent REBL-PAT hPSCs were enzymatically dissociated, washed in PBS and fixed for 2 hrs in 70 % ethanol, washed again in PBS, stained with propidium iodide solution (PBS solution supplemented with 10 μg/ml propidium iodide, 0.1 % Triton X-100 and 100 μg/ml RNase A) and sorted based on the PI staining intensity into G0/G1, S and G2/M cell cycle phases at 98% purity after excluding debris and apoptotic cells with forward vs side scatter gating using Beckman Coulter Astrios EQ and Beckman Coulter Kaluza Analysis 2.1 software.

### 2.2.4. Cell culture for cancer cell lines

#### 2.2.4.1. LN18, LN229, and U87MG glioblastoma cell lines

LN18, LN229, and U87MG were cultured in DMEM (high glucose) supplemented with 10% heat-inactivated fetal bovine serum and 1% penicillin/streptomycin at 37 °C in a humidified incubator with 5% CO2. The culture medium was replaced daily. To passage cells (every 3-4 days), the culture medium was discarded, and the cells were washed with sterile PBS, before dissociating them with trypsin for 3 minutes at 37 °C with 5% CO2. The excess trypsin in dissociated cells was quenched with culture medium and the cells were pelleted by centrifugation for 4 min at 160g. After discarding the supernatant, the cells were resuspended in complete medium and incubated at 37 °C in a humidified incubator with 5% CO2 until next split or collected for experimentation.

#### 2.2.4.2. HeLa cell culture

HeLa cells were cultured essentially in the same conditions as in 2.2.3.2. The principal difference was that HeLa cells were passaged every 3$^{rd}$ day. The culture was replaced with a frozen batch of low passage cells after 20 passages.

#### 2.2.4.3. HAP1 cell culture

Wild type and YTHDF2 KO HAP1 cells were cultured on DMEM/F12 (Gibco Life Technologies, no. 11320033) supplemented with 20% heat-inactivated fetal bovine serum containing 1% pen/strep at 37 °C in a humidified incubator with 5% CO2. Cells were passaged when reached between 70-80 % confluency. Cells were regularly inspected to confirm the largely haploid genome is maintained and the culture is replaced with a frozen batch of low passage cells after 10 passages.

### 2.2.4. Wild Type and *Ythdf2* KO mouse Neural Stem cell culture

Wild Type and Ythdf2 KO mouse Neural Stem cells were cultured in our collaborator's lab as described in (Li et al., 2018).

### 2.2.5. Growth curve of Wild type and YTHDF2 KO HAP1 cells

After reaching about 60 % confluency, wild type and YTHDF2 KO HAP1 cell cultures were passaged and seeded 4x T25 tissue culture flasks at $350x10^4$ cell each, this was labelled day zero. After 24hrs, 1x T25 flask from WT and YTHDF KO cells were dissociated and the total number was determined using hemocytometer. This was repeated every 24 hrs for 96 hrs in total. The same cell counting was repeated twice more using a fresh batch of WT and YTHDF2 KO cells in each case. The average number of both WT and YTHDF2 KO for each day was plotted against time (hrs). Statistical significance was determined using two-tailed t-test following assessment of variance with F-test.

### 2.2.6. Short interference RNA knockdown of target genes

For depletion of target transcripts (*METTL3*, *YTHDF2*, *HNRNPA2B1*, *RNase H1*), hPSCs were transfected with 50 pmol of the corresponding siRNA duplex using DharmaFECT (GE Lifesciences) in antibiotic free medium (Optimum), according to Manufacturer's instructions. After 24 hrs of incubation at 37 °C in a humidified incubator with 5% CO2, the medium was repeated once more and the cells incubated for further 24 hrs at 37 °C in a humidified incubator with 5% CO2. The next day, the cell were harvested for either RT-qPCR, DRIP, ChIP or for staining if transfection was performed on hPSCs cultured in Matrigel-coated 8-well glass chamber slides. As a control, hPSCs were transfected with a non-targeting siRNA.

### 2.2.7. pEGFP-RNASEH1 plasmid transfections in HAP1 cells

For overexpression of RNase H1, 2 µg of a C-terminally eGFP-tagged human RNASEH1 (nuclear isoform) pEGFP-RNASEH1 plasmid was transfected into wild type and YTHDF2 KO HAP1 cells using the FuGENE® HD Transfection Reagent (Promega) at 3:1 ratio as per the

manufacturer's instruction. A parallel transfection with the pmaxGFP (Lonza) was used as a control green fluorescent protein (GFP)-only plasmid. The cells were incubated at 37 °C in a humidified incubator with 5% CO2. The transfection was repeated after 24hrs and the cells were harvested for staining or DRIP.

### 2.3. Molecular Biology Experiments

#### 2.3.1. Growth of pEGFP-RNASEH1 plasmid containing bacteria

To amplify the pEGFP-RNASEH1 plasmid DNA for transfections, the stab of pEGFP-RNASEH1 plasmid containing bacteria was spread on Kanamycin (50 µg/mL) containing LB-agar plates and a single colony from the surviving clones was grown overnight at 37 °C in 10 mL of Kanamycin (50 µg/mL) containing LB medium.

#### 2.3.2. Isolation of pEGFP-RNASEH1 plasmid

Plasmid isolation was performed using Qiagen's Qiaprep Spin Miniprep kit as instructed by the manufacturer. The purified plasmid was quantified prior to use verifications.

#### 2.3.3. Verification of pEGFP-RNASEH1 plasmid

Following isolation, pEGFP-RNASEH1 plasmid was verified by sequencing.

#### 2.3.4. Verification of Wild type and YTHDF2 KO HAP1 cells.

YTHDF2 KO HAP1 cells was generated by the manufacturer via CRISPR/Cas9-mediated deletion of 140bp in exon 3, leading to frameshift and generation of premature stop codon (Figure 2.3A). The deletion was confirmed by sequencing (Figure 2.3B) and PCR (D) with primers highlighted in B.

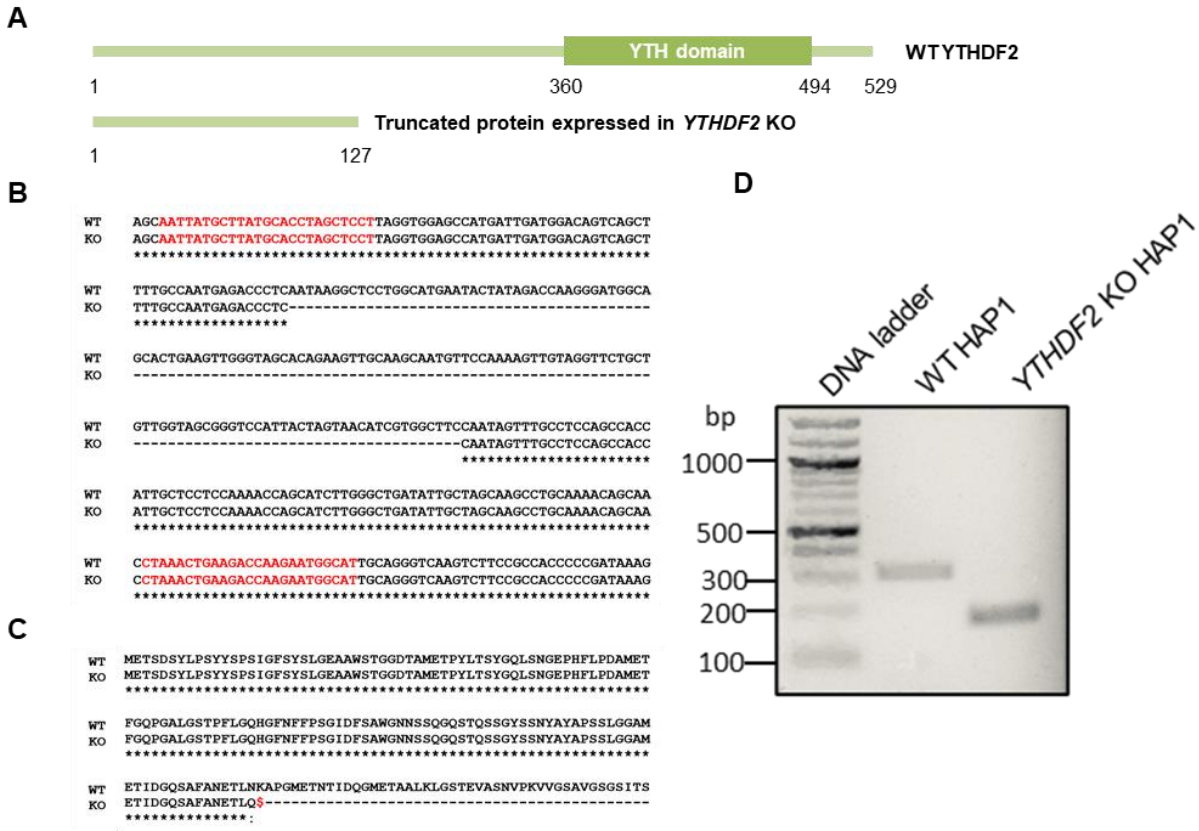



**Figure 2.4. Validation of YTHDF2 KO in HAP1 cells**. (A) Schematic illustrating structure of wild type (WT) YTHDF2 protein and truncated protein expressed in YTHDF2 KO HAP1 cells. The lengths (aa) of the proteins and YTH domain are shown. (B) The alignment of genomic sequences of YTHDF2 exon 3 in WT and YTHDF2 KO (KO) HAP1 cells. (C) The alignment of protein sequences corresponding to YTHDF2 exon 3 in WT and YTHDF2 KO (KO) HAP1 cells. The premature stop codon generated by deletion/frameshift in KO is presented in red. (D) The results of PCR performed using genomic DNA isolated from WT and YTHDF2 KO HAP1 cells and the primers corresponding to the sequences of YTHDF2 exon 3 are depicted in red in (B).

### 2.3.5. Functional validation of YTHDF2 KO in HAP1 cells by immunostaining.

The 140 bp deletion from the c-terminal region of YTHDF2 impacts its YTH-(m6A binding)-domain. Therefore, it was hypothesised in YTHDF2 KO HAP1, if any, only a truncated version would exist resulting from early stop codon. This was verified by immunostaining wild type and YTHDF2 KO HAP1 cells for YTHDF2 as described in 2.4.8. In contrast to wild type cells in which YTHDF2 staining was detected in the mitotic chromatin, YTHDF2 KO cells did not exhibit mitotic YTHDF2 staining. These results further confirmed the functional impairment the 140 bp deletion causes on YTHDF2.

## 2.4. Immunostaining experiments

### 2.4.1 m6A immunocytochemistry in non-denaturing conditions

For m6A staining in non-denaturing (no HCl treatment) conditions, to each well of a 8-well Matrigel coated chamber slide, approximately $13x10^4$ REBL-PAT hiPSCs were seeded, and the cells were fixed for 15 minutes in 4 % paraformaldehyde (PFA) after reaching 60-70% confluency. After safely discarding the fixative, the cells were washed 3 times, for 5 minutes each, with PBS, permeabilized with PBS solution containing 0.5 % Triton X-100 and washed 5 times for 5 minutes each with PBS to remove excess detergent. Following permeablization, the cells were blocked in PBS containing 10 % FBS before replacing the blocking solution with 1:200 dilution of rabbit anti-m6A antibody in blocking buffer. After 30 minutes incubation at room temperature in a humidified chamber, excess antibody was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01 tween 20 before applying goat anti-rabbit HRP-conjugated secondary antibody (1:400 dilution) in blocking solution for an hour in humidified chamber at room temperature. Excess secondary antibody was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 followed by 2 minutes incubation in Tyramide (1:200 TSA buffer) solution. The slide was washed 3 times for 5 minutes each before coating the slide with DAPI containing mounting medium and sealed for examination using microscopy. In parallel, slides with no primary and no secondary antibodies were included as controls.

### 2.4.2. Co-immunostaining of m6A and 5-mC in denaturing conditions

For m6A staining in denaturing (with HCl treatment) conditions, to each well of a 8-well chamber slide, approximately $13x10^4$ REBL-PAT hiPSCs were seeded, and the cells were fixed for 15 minutes in 4 % paraformaldehyde (PFA) after reaching 60-70 % confluency. After safely discarding the fixative, the slide containing the cells was washed 3 times, for 5 minutes each,

with PBS, permeabilized with PBS solution containing 0.5 % Triton X-100 and washed 5 times for 5 minutes each with PBS to remove excess detergent. Following permeablization, the slides containing the cells were incubated for 1 hr at room temperature in 4N HCL solution. Following the DNA denaturation step, the 4N HCl was safely discarded and the slides were washed 5 times, 5 minutes each, with PBS to remove residual HCl. Next, the slides were blocked in PBS containing 10 % FBS before replacing the blocking solution with a mixture of 1:200 dilution of rabbit anti-m6A antibody and 1:100 dilution of mouse anti-5mC antibody in blocking buffer. After 1 hr incubation in a humidified chamber at room temperature, excess antibody was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 before incubating in blocking buffer containing a mixture of goat anti-rabbit HRP-conjugated secondary antibody (1:400 dilution) and 1:400 dilution of Alexafluor 633-conjugated secondary antibody for an hour in humidified chamber at room temperature. Excess secondary antibody mixture was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 followed by 2 minutes incubation in Tyramide (1:200 TSA buffer) solution. The slide was washed 3 times for 5 minutes each before coating the slide with DAPI containing mounting medium and sealed for examination using microscopy. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.3. Co-immunostaining of m6A and PH-3, 6mA and 5mC in no RNase condition, RNase A only and RNases A and H.

The staining was essentially performed using previously described procedure (Abakir et al., 2016). Briefly, for m6A co-staining with PH-3 in denaturing (with HCl treatment) conditions, each well of a 8-well Matrigel-coated chamber slide was seeded with approximately $13x10^4$ REBL-PAT hiPSCs, and the cells were fixed for 15 minutes in 4% paraformaldehyde (PFA) after reaching 60-70% confluency. After safely discarding the fixative, the slide containing the

cells was washed 3 times for 5 minutes each with PBS, permeabilized with PBS solution containing 0.5% Triton X-100 and washed 5 times, for 5 minutes each with PBS to remove excess detergent. Following permeablization, the slides were separates into three groups: No RNases treated, treatment with only 25 mg/mL RNase A, or treatment with RNase H buffer containing a mixture of 25 mg/mL RNase A and 10U of RNase H and incubated overnight at 37 °C. After discarding excess RNase solution, the slides were then incubated 1hr at room temperature in 4N HCL solution for DNA depurination. Following the DNA denaturation step, 4N HCl was safely discarded and the slides were washed 5 times, for 5 minutes each, with PBS to remove residual HCl. The slides were next blocked in PBS containing 10% FBS before replacing the blocking solution with a mixture of 1:200 dilution of rabbit anti-m6A antibody and 1:200 dilution of mouse anti-phospho-Histone H3 antibody or 1:200 dilution of rabbit anti-m6A antibody and 1:100 dilution of mouse anti-5mC antibody in blocking buffer. After 1 h incubation in a humidified chamber at room temperature, excess antibody mixtures were removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 before incubating in blocking buffer containing a mixture of goat anti-rabbit HRP-conjugated secondary antibody (1:400 dilution) and Alexafluor 633-conjugated secondary antibody for an hour in humidified chamber at room temperature. Excess secondary antibody mixture was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 followed by 2 minutes incubation in Tyramide (1:200 TSA buffer) solution. The slide was washed 3 times for 5 minutes each before coating the slide with DAPI containing mounting medium and sealed for examination using microscopy. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.4. Co-immunostaining of YTHDF2 and RNA:DNA hybrids (S9.6)

The staining was performed using previously described procedure (Sollier et al., 2014). Briefly, for YTHDF2 co-staining with S9.6, each well of 8-well Matrigel-coated chamber slide was seeded with approximately $13x10^4$ REBL-PAT hiPSCs, and after reaching 60-70% confluency, cells were fixed for 15 minutes in 4% paraformaldehyde (PFA). After safely discarding the fixative, the cells were washed 3 times for 5 minutes each, with PBS, permeabilized with PBS solution containing 0.5% Triton X-100 and washed 5 times, for 5 minutes each, with PBS to remove excess detergent. Next, the slides were blocked in PBS containing 10% FBS for 1 h before replacing the blocking solution with a mixture of 1:500 dilution of rabbit anti-YTHDF2 antibody and 1:200 dilution of mouse anti-phospho-Histone H3 antibody or 1:200 dilution of rabbit anti-m6A antibody and a 1:500 dilution of mouse S9.6 monoclonal antibody in blocking buffer. After 1 h incubation in a humidified chamber at room temperature, excess antibody mixtures were removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 before incubating in blocking buffer containing a mixture of goat anti-rabbit HRP-conjugated secondary antibody (1:400 dilution) and Alexafluor 633-conjugated secondary antibody for an hour in humidified chamber at room temperature. Excess secondary antibody mixture was removed with 3 washes, 5 minutes each, using PBS solution containing 0.01% tween 20 followed by 2 minutes incubation in Tyramide (1:200 TSA buffer) solution. The slides were washed 3 times for 5 minutes each before coating the slide with DAPI containing mounting medium and sealed for examination using microscopy. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.5. Co-immunostaining of YTHDF1 and phosphorylated histone H3 in hPSCs cells

The same immunostaining procedure as in 2.4.4 was followed except for primary antibodies. Blocking solution contained a mixture of a 1:500 dilution of rabbit anti-YTHDF1 and 1:200 dilution of mouse anti-phospho-Histone H3 antibody. In parallel, slides with no primary and

no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.6. Co-immunostaining of YTHDF2 and phosphorylated histone H3 in hPSCs.

The same immunostaining procedure as described in 2.4.4 was followed except for primary antibodies. Blocking solution contained a mixture of 1:500 dilution of rabbit anti-YTHDF2 antibody and 1:200 dilution of mouse anti-phospho-Histone H3 antibody. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.7. Co-immunostaining of HNRNPA2B1 and m6A in hPSCs

The immunostaining procedure described in 2.4.4 was followed except for primary antibodies. Blocking solution contained a mixture of 1:200 dilution of mouse anti-HNRNPA2B1 antibody and 1:200 dilution of rabbit anti-m6A antibody. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.8. Co-immunostaining of YTHDF2 and phosphorilated histone H3 in WT and YTHDF2 KO cells

The same immunostaining procedure as in 2.4.4 was followed except that HAP1 cells were seeded onto attachment factor protein coated 8-well chamber slides and, for primary antibodies, blocking solution containing a mixture of 1:500 dilution of rabbit anti-YTHDF2 antibody and 1:200 dilution of mouse anti-phospho-Histone H3 antibody was used. In parallel, slides with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.4.9. Immuno-competition assay for validation of specificity of anti-m6A antibody

To examine specificity of the anti-m6A antibody, hPSCs were co-stained as described in 2.4.2 for m6A and 5-hydroxymethylcytosine (5hmC) in the presence of no competitor, N6-methyl-2′-deoxyadenosine-5′-triphosphate (6mATP), unmodified deoxyadenosine triphosphate (ATP) or deoxythymidine triphosphate (TTP) from deoxynucleotide triphosphate set.

### 2.4. 10. Immunostaining of 1-methyladenosine (m1A) in hPSCs

The m1A staining was performed as described in 2.4.1 except the anti-m6A antibody was replaced with 1:500 dilution of mouse anti-m1A monoclonal antibody in blocking buffer.

### 2.4.11. Immunostaining of METTL3 KD hPSCs for phosphorylated γH2AX

HPSCs cultured in Matrigel-coated 8-well chamber slides were depleted from *METTL3* as described in 2.2.7 along with control and stained for phosphorylated γH2AX as described in 2.4.2 except anti-m6A antibody was replaced with 1:200 dilution of mouse anti-γH2AX antibody in blocking buffer.

### 2.4.12. Immunostaining of HNRNPA2B1 KD hPSCs for phosphorylated γH2AX

hPSCs cultured in Matrigel-coated 8-well chamber slides were depleted from *METTL3* as described in 2.2.7 along with controls and stained for phosphorylated γH2AX as described in 2.4.2 except anti-m6A antibody was replaced with 1:200 dilution of mouse anti-γH2AX antibody in blocking buffer.

### 2.4.13. Immunostaining of phosphorylated γH2AX in WT and YTHDF2 KO HAP1 cells

HAP1 WT and KO cells cultured in 8-well chamber slides were stained for phosphorylated γH2AX as described in 2.4.2 except anti-m6A antibody was replaced with 1:200 dilution of mouse anti- γH2AX antibody in blocking buffer.

### 2.4.14. Immunostaining of wild type and Ythdf2 KO mouse embryonic brain sections for phosphorylated γH2AX.

Generation of *Ythdf2* conditional KO mice, followed by cre-mediated deletion and derivation of mNSCs from E14.5 embryonic forebrains, was described by Li and colleagues (Li et al., 2018). All mouse experiments were approved by the Norwegian Animal Research Authority, Norwegian Food Safety Authority and performed in accordance with institutional guidelines at the Center for Comparative Medicine at Oslo University Hospital. Animal work was conducted in accordance with the rules and regulations of the Federation of European Laboratory Animal Science Association (FELASA). Immunohistochemistry was essentially performed as previously described in (Abakir et al., 2016). Paraffin-embedded E14.5 WT and *Ythdf2* KO mouse embryonic brain sections were first dewaxed using 2 washes of 10 minutes each in xylene at room temperature. Following dewaxing, the tissues were rehydrated by successive 10-minute washes in 95, 75, and 50% ethanol at RT. Rehydrated tissues were fixed in 4% paraformaldehyde (PFA) at room temperature. Following 15 minutes of incubation, excess fixative was safely discarded, and the sections were washed 3 times for 5 minutes each in PBS before permeabilization in PBS containing 0.5 % Triton-X100 for 30 minutes at room temperature. Next, excess detergent was removed by 3 washes of 5 minutes each prior to incubation in blocking buffer (PBS containing 10% FBS) for 1 h at room temperature. The blocking solution was replaced with 1:200 dilution of mouse anti- γH2AX primary antibody in blocking buffer. Following 1 h incubation at room temperature, the unbound primary antibody was removed with 3 washes of 5 minutes 0.01% Tween 20 in PBS before incubation in 1:400 dilution of Alexafluor 633-conjugated secondary antibody for an hour in humidified chamber at room temperature. Excess secondary antibody was removed with 3 washes, 5 minutes each using PBS solution containing 0.01% tween 20. The sections were washed 3 times for 5 minutes each before coating the slide with DAPI containing

mounting medium and sealed for examination using microscopy. In parallel, sections with no primary and no secondary antibodies were included as controls. This staining was repeated three times using biologically independent samples.

### 2.5. Microscopy and image analysis

#### 2.5.1. Widefield microscopy

Nikon Eclipse 90i fluorescent microscope was employed to rapidly examine whether staining experiment worked prior to imaging the samples using confocal microscope.

#### 2.5.2. Confocal microscopy images

Representative images (500-nm optical sections) of immunostained mammalian cells were acquired with a Zeiss LSM 780 confocal microscope using a Plan Apochromat ×63/1.40 numerical aperture Oil DIC M27 objective. The images for each experiment were obtained using identical acquisition protocol, exposure and gain settings to ensure appropriate comparisons between samples.

#### 2.5.3. Imaging analysis

Confocal images were processed using Image J (Schneider *et al.*, 2012) and Adobe Photoshop. Unless stated otherwise, from each representative image, a minimum of 20 nuclei were selected for signal intensity quantifications using the ZEN Zeiss LSM 780 imaging software as described in (Ramsawhook et al., 2017). Statistical significance was determined GraphPad Prism using 2-tailed t-test following assessment of the variance with F.

### 2.6. Liquid Chromatography–Tandem Mass Spectrometry (LC–MS/MS)

#### 2.6.1. Preparation of cells for DNA and RNA isolation

To prepare genomic and total RNA for assessment of 6mA and m6A using LC-MS/MS, 70-80% confluent mitotic cell-enriched (KaryoMax-colcemid) and asynchronous control REBL-

PAT hiPSCs and HUES7 hESCs cultures were washed in PBS, dissociated by TrypLE Select Enzyme, washed with PBS again and pelleted into 1.5 mL Eppendorf tubes following inactivation of TrypLE Select Enzyme with BGK solution.

**2.6.2. Isolation of genomic DNA and RNA.**

Genomic DNA and total RNA was isolated from pelleted hPSCs (REBL-PAT hiPSCs and HUES7 hESCs) using DNeasy Blood and Tissue Kit and RNeasy mini Kit respectively following manufacturer's instruction. Approximately 200 µg of DNA and total RNA samples were isolated and kept in -20 °C freezer until examined by our collaborators in NEB for m6A/6mA presence and quantification.

**2.6.3. Detection and quantification of m6A in genomic DNA and total RNA using LC–MS/MS**

These experiments were performed by our collaborators (Dr Daniel Gackowski's group) in Department of Clinical Biochemistry, Nicolaus Copernicus University in Toruń, Bydgoszcz, Poland. Approximately 15 µg of purified DNA and RNA were used for LC–MS analysis. Genomic DNAs and total RNAs were digested to nucleosides by treatment with Nucleoside Digestion Mix (NEB, M0649S) overnight at 37 °C. LC–MS/MS analysis was performed in duplicate by injecting digested DNAs and RNAs on an Agilent 1290 UHPLC equipped with a G4212A diode array detector and a 6490A Triple Quadrupole Mass Detector operating in positive electrospray ionization mode. UHPLC was carried out using a Waters XSelect HSS T3 XP column (2.1×100mm2, 2.5 µm) with the gradient mobile phase consisting of methanol and 10mM aqueous ammonium formate (pH 4.4). MS data acquisition was performed in dynamic multiple reaction monitoring mode. Each nucleoside was identified in the extracted chromatogram associated with its specific MS/MS transition: dC at m/z 228→112, d5mC at m/z 242→126, d5hmC at m/z 258→142, d5fC at m/z 256→140, dA at 252→136, d6mA at 266→150, rC at m/z 244→112, m5 C at 258→126, Cm at 258→112, rA at 268→136, m1 A

at 282→150, Am at 282→136, m6 A at 282→150 and m6 2A at 296→164. External calibration curves with known amounts of the corresponding nucleosides were used to calculate ratios within the samples analysed. Two independent biological replicates were performed for analysis of m6A in genomic DNA and three biological replicates for m6A assessment in total RNA extracted from human cells.

### 2.7. Genome-wide mapping of m6A modified RNA:DNA hybrids in mammalian cells

#### 2.7.1. DNA isolation

Genomic DNA was isolated from hPSCs (REBL-PAT hiPSCs and HUES7 hESCs) by SDS/Proteinase K treatment at 37 °C followed by incubation with 100 µg/ml$^{-1}$ RNaseA for 30 min in lysis buffer, phenol-chloroform extraction, and ethanol precipitation.

#### 2.7.2. Quantification of genomic DNA.

Quick assessment of DNA purity and quantification was initially performed using NanoDrop™, then all subsequent genomic DNA and RNA quantifications were performed using Qubit™ as instructed by the manufacturer.

#### 2.7.3. Sonication of genomic DNA

The DNA was fragmented to 300–600bp using Covaris S2 ultrasonicator (Covaris). Sonication was confirmed using the Agilent 2200 TapeStation system (Figure 2.7.3)

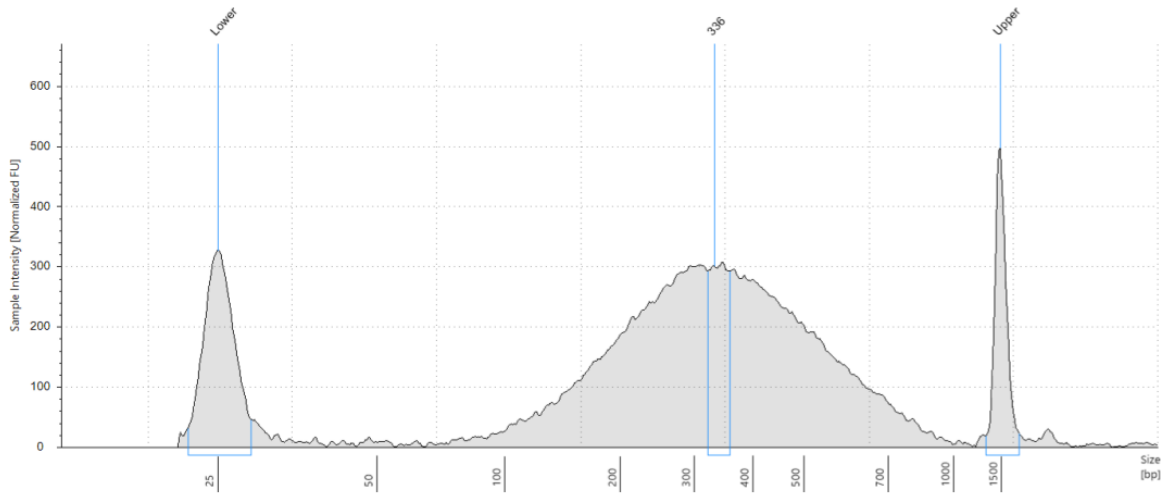



**Figure 2.7.3. Visualization of a typical sample of sonicated DNA fragments using Agilent 2200 TapeStation system and D1000 reagent.**

### 2.7.4. RNase H treatment of control samples for m6A and S9.6 DRIP

Following sonication, genomic DNA was divided in two fractions, one treated with RNase H treatment and one without such treatment. RNase H treatment was performed with 10 units of RNase H (NEB, no. M0297S) in 1×RNase H buffer overnight at 37 °C. Next day, the samples were passed through DNA extraction columns (DNeasy Blood & Tissue Kit) prior to adaptor ligation.

### 2.7.5. DNA adaptor ligation and library preparation for Illumina sequencing

Both RNase H treated and untreated genomic DNA samples were ligated with NEB adaptors using the NEBNext® Multiplex Oligonucleotides kit as instructed by the manufacturer's protocol.

### 2.7.6. S9.6 immunoprecipitation

S9.6 DRIP was carried out essentially as described previously by the Chedin lab (Ginno et al., 2012). For every DRIP reaction (RNase H treated and untreated), 10 µg of sonicated genomic DNA was diluted in 400 µL of nuclease free water before adding 50 µL of 10x IP buffer and

10 µL of S9.6 antibody. The mixture was incubated overnight in an orbital shaker at 4 °C. To the mixture, 20 µl of PBS washed Sheep anti-mouse IgG (Dynabeads) were added the next day and incubated for 2 hrs in an orbital shaker at 4 °C. After 2 hrs, the samples were placed with lids open on a magnetic rack and the supernatant (containing non-precipitated fragments) was replaced with two 1xIP buffer washes, 10 minutes each, before returning the samples on the magnetic rack to remove the excess IP buffer. The samples containing R-loop enriched fragments were then incubated overnight at 55 °C in proteinase K buffer containing 3 µL of 25mg/mL proteinase K to digest the antibody. The nucleic acids were isolated by column purification using QIAquick PCR purification kit. The corresponding primary IgG-only and secondary IgG-only (Dynabeads-only) DRIP reactions were used in control immunoprecipitations.

### 2.7.7. m6A RNA:DNA immunoprecipitation (m6A DIP)

In order to map m6A-modified RNA:DNA hybrids using the anti-m6A antibody, previously described methylated DNA immunoprecipitation (meDIP) approach (Weber et al., 2005) was adapted. In every m6A DIP reaction (RNase H treated and untreated), 10 µg of sonicated genomic DNA was diluted in 400 µL of nuclease free water and denatured for 10 min at 95 °C prior to incubation with 50 µL of 10x IP buffer and 10 µL of anti-m6A antibody. The mixture was incubated overnight in an orbital shaker at 4 °C. To the mixture, 20 µl of PBS washed sheep anti-rabbit IgG (Dynabeads) were added the next day before incubation for 2 hrs in an orbital shaker at 4 °C. After 2 hrs, the samples were placed with lids open on a magnetic rack and the supernatant (containing non-precipitated fragments) was replaced with two 1x IP buffer washes, 10 minutes each, before returning the samples on the manganic rack to remove excess IP buffer. The samples containing R-loop enriched fragments were then incubated overnight at 55 °C in proteinase K buffer containing 3 µL of 25 mg/mL proteinase K to digest the antibody. R-loops were isolated by column purification using QIAquick PCR purification kit.

The corresponding primary IgG-only and secondary IgG-only (Dynabeads-only) DRIP reactions were used in control immunoprecipitations.

**2.7.8. Library preparation and high throughput m6A DIP and S9.6 DRIP sequencing**.

Following immunoprecipitation, adapter-ligated RNA:DNA hybrid-enriched samples were amplified for 15 cycles as described in NEB Next DNA Ultra Library Preparation Kit. The libraries were quantified using the Kapa Library Quantification Kit. High-throughput sequencing was performed using Illumina NextSeq500 platform to generate 2×150-bp reads. Primary IgG-only DRIP reactions resulted in levels of DNA insufficient for successful library production, even with the maximum number (15) of amplification cycles recommended by NEB. A summary of both m6A DIP and S9.6 DRIP is outlined below (Figure 2.7.8).

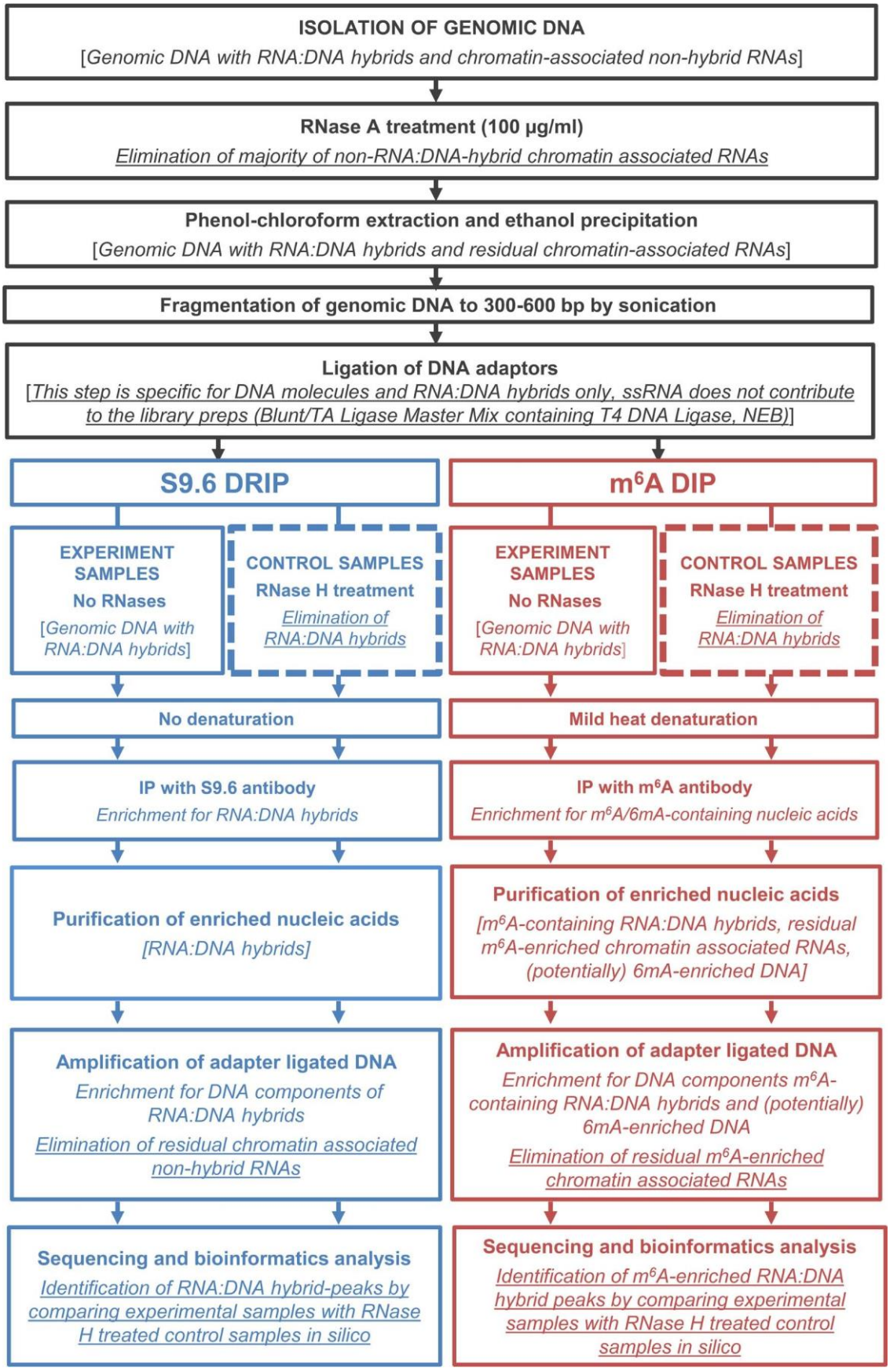



**Figure 2.7.8. Workflow for m6A DIP and S9.6 DRIP sequencing techniques.**

### 2.8. Whole-transcriptome sequencing

#### 2.8.1. Total RNA extraction

RNA extraction from hPSCs was performed using the RNeasy mini kit (QIAGEN) following manufacturer's instructions. Importantly, on-column 15 minutes at room temperature DNase digestion using the RNase-Free DNase Set was performed to remove any residual contaminating DNA.

#### 2.8.2. RNA library preparation and sequencing

RNA sequencing libraries were constructed using Illumina TruSeq Stranded Total RNA sample preparation kit according to the manufacturer's guidelines, and then sequenced on an Illumina HiSeq 4000 generating 20–50 million 75-bp paired-end reads per sample.

### 2.9. Bioinformatic analysis

Bioinformatic analysis was carried out in collaboration with the University of Nottingham ADAC team.

#### 2.9.1. Analysis of m6A DIP- and S9.6 -seq data

The 150-bp Illumina paired-end reads were trimmed using Skewer to remove low-quality sequences (Jiang et al., 2014). Reads that passed filtering were aligned to the human Ensembl genome (build hg38.89) using the Burrows–Wheeler aligner with default parameters (Li and Durbin, 2009). As each biological sample was split across multiple lanes of sequencing, the corresponding alignments were merged with Samtools (Li et al., 2009) and de-duplicated to remove PCR artifacts with picard-tools MarkDuplicates (Chang et al., 2013). The impact of each pulldown was assessed using Phantompeakqualtools (Marinov et al., 2014), and highly modified regions (peaks) were identified using MACS2.1.1 (Zhang et al., 2008; Marinov et al., 2014). Because the exact mode of genomic distribution of m6 A-containing RNA:DNA hybrids was initially unknown, detection of both narrow and broad peaks was carried out using -q0.01 settings for narrow peaks and -broad-cutoff 0.1 (q0.01) for broad peaks. High-confidence peaks

and consensus peaks were identified using the bioconductor package DiffBind (Ross-Innes et al., 2012). We performed peak calling against input DNA and against secondary IgG-only control samples. More than 96% of the m6A DIP peaks called against input were also identified using IgG-only controls. Peaks called against input were used for further analysis. Consensus peaks were defined using the dba.peakset () function to select for peaks overlapping in both replicates. In each instance the replicate sample BAM/bed files, along with the corresponding input samples, were used as input. The R package ChipQC (Carroll et al., 2014) was employed for assessing the peaks' genomic distribution and the quality of the sequencing data. The Ensembl Genome GTF file (Homo_sapiens.GRCh38.89.gtf) (Zerbino et al., 2018) was downloaded and processed to extract gene biotypes, annotated gene features, introns, intergenic regions and sequences 3 kb up/downstream of genes. The Repeatmasker Library (db20140131) (Bao et al., 2015) was downloaded from UCSC, converted to the ensembl gene build and processed to identify repeat masked regions. The Ensembl regulatory build (HgGRCh38.Regulatory_Build.regulatory_features.20161111) was downloaded and processed to identify regulatory features (Zerbino et al., 2015). The consensus and DMR bedfiles were annotated with this list of gene features, gene biotypes, repeat-masked sequences and regulatory regions using BedTools intersect (Quinlan and Hall, 2010). The Features present in each of the samples were then counted using awk and processed into a matrix.

**2.9.2. Bioinformatics: Comparison of the m6A DIP and S9.6 DRIP peaks datasets with genomic distribution of histone modifications in hPSCs.**

Previously released hPSCs ChIP– seq datasets for histone marks (H3K4me1, GSM733782; H3K4me2, GSM733670; H3K4me3, GSM733657; H3K27me3, GSM733748; H3K27ac, GSM733718; H3K9me3, GSM1003585; H3K36me3, GSM733725; H3K79me2, GSM1003547) were compared against consensus peaks using an adaptation of Aaron Quinlans bedtools jacard statistic pipeline (https://github.com/arq5x/tutorials/blob/master/bedtools.md).

### 2.9.3. Analysis of gene expression for the m6A and S9.6 peaks-containing genes

The whole transcriptome sequencing raw reads were trimmed using Skewer and processed using the standard Tuxedo v2 pipeline; reads were aligned to the human genome hg38.89 (from biomart) using Hisat2 (Frazee et al., 2015), then novel transcripts were identified and quantified using Stringtie. Significantly differentially expressed genes were identified using the default Ballgown parametric F-test (Hulsen et al., 2008). The mean levels of expression of genes with/without peaks (both across the entire length of genes +/- 3 kb and restricted to peaks within specific feature types) were compared and a series of graphs generated using R. The list of hPSCs promoter categories defined in the previous study (Verma et al., 2018) was kindly provided by Danwei Huangfu (Memorial Sloan Kettering Cancer Center). The promoters were characterised in the previous study as active if they were associated with H3K4me3 (in the 1.5-kb region flanking the TSS) and H3K79me2 (in the 5-kb region downstream of the TSS), as initiated if they were associated with H3K4me3 (in the 1.5-kb region flanking the TSS), as bivalent if they were associated with H3K4me3 and H3K27me3 (in the 1.5 kb region flanking the TSS) and as silent if they were not associated with H3K4me3 (in the 1.5 kb region flanking the TSS) (Verma et al., 2018). Area-proportional Venn diagrams for lists of peaks/genes were generated using BioVenn (Hulsen et al., 2008). For analysis of distribution of DRIP and control reads in the areas containing m6A only-, S9.6 only- and m6A/S9.6 overlapping peak sets, the peaks called with MACS2 were compared against replicate-merged bamfiles using Deeptools compute matrix tool to create a per peak-set distribution matrix of read coverage +/- 3 Kb from the center of the peak for each biological condition (Ramírez et al., 2016). The Deeptools plotHeatmap - averageTypeSummaryPlot tool was used to visualise the results. For comparison of gene expression against the number of m6A and S9.6 DRIP peaks per gene, the average per gene FPKM values were extracted from the RNAseq data and directly compared against the

number of peaks per gene (including +/- 3 Kb flanking regions) for the m6A and S9.6 diffbind consensus peak-sets. Scaled binning was employed to dimensionally reduce the scope of the data in order to effectively visualize it.

### 2.9.4. The distribution of m6A marked RNA:DNA hybrids among different repetitive sequences.

For LINE1, Alu, ERV and SVA enrichment analyses, raw Illumina RNAseq reads from H7-hESCs (GSE36552) (Yan et al., 2013) and raw Illumina reads from m6A DIP and S9.6 DRIP (2 replicas each) were aligned to the consensus sequences of L1Hs-Ta, L1Hs-preTa, L1PA2, L1PA5, AluY, AluS, SVA-D, SVA-E, SVA-F, HERV-LTR5_Hs and HERV-LTR7Y extracted from Repbase (http://www.girinst.org/repbase/) (Bao *et al*., 2015). Alignment used Bowtie2 (Langmead et al., 2009) using end-to-end and sensitive parameters: -D 15 -R 2 -N 0 -L 22 -i S, 1, 1.15. Mapped reads were counted and normalized by retroelement length in kb and by the total number of reads mapped to the genome in millions (RPKM, Reads Per Kilobase per Million mapped reads). The annotated bed files have been deposited to the following online repository (https://bitbucket.org/ADAC_UoN/adac1075-bed-files/). The in-house scripts used for the analysis can be found in the following online repository (https://bitbucket.org/ADAC_UoN/adac0175-code/). These analyses were performed in an excellent collaboration with the Garcia-Perez lab in the MRC Human Genetics Unit together ADAC team in Nottingham.

## 2.10. Validation of m6A DIP and S9.6 DRIP protocols

### 2.10.1. Two-round RNA:DNA hybrid immunoprecipitations (S9.6 DRIP followed by m6A DRIP)

For the two-round (S9.6 DRIP followed by m6A DRIP) DRIP/DRIP, 10 individual S9.6 DRIP reactions were performed in parallel that generated approximately 500 ng of total nucleic acid. The recovered nucleic acid was used for m6A DIP followed by qPCR analysis of R-loops

enriched intronic and repetitive sequences. This approach demonstrated that both m6A and S9.6 recognised essentially the same RNA:DNA hybrids, thus validating the m6A DRIP procedure.

**2.10.2. Validation of m6A DRIP procedure: m6A RIP qRT-PCR following S9.6 DRIP**

For S9.6 DRIP followed by the m6 A RIP experiment, nucleic acids recovered from DRIP were denatured for 30 min at 95 °C. This was followed by digestion of the DNA components of RNA:DNA hybrids using Turbo DNase for 30min at 37 °C. After inactivation of DNase, approximately 100 ng of the recovered RNA was used for m6A RIP performed using the EpiMark N6-Methyladenosine Enrichment Kit. The recovered RNA was reverse transcribed using SuperScript III Reverse Transcriptase and random hexamers and analyzed by qPCR. IgG-only reactions and reactions carried out without reverse transcription were used as controls.

**2.10.3. Validation of m6A DIP and S9.6 DRIP using synthetic oligoes.**

For validation of m6A DIP using synthetic oligonucleotides, 0.1–1.0 pmol of m6A-containing and non-modified RNA:DNA hybrids and individual single stranded RNA or DNA oligonucleotides were spiked-in with 5 µg of mouse gDNA. To generate RNA:DNA hybrids, RNA and DNA oligonucleotides were mixed in equimolar concentrations, incubated for 20 min at 98 °C, slowly cooled in a heating block and placed on ice. Quantitative PCR analysis of m6A DIP and S9.6 DRIP samples was carried out with SYBR Green PCR Master Mix (Sigma) according to standard procedures. Fold enrichment was calculated as 2ddCt, where dCt=Ct(enriched)–Ct(input) and ddCt=dCt–Ct (IGG). The primers used for DRIP/DIP–qPCR and qPCR analysis of the corresponding transcripts are listed in Supplementary Table 1. Primers for α-satellites were obtained from Novus (no. NBP1-71654SS). Generic primers amplifying Alu elements from the indicated families and evolutionarily young L1Hs were used. Human LINE1 primers were designed to detect L1PA1 and L1PA2 classes of these retro-elements. Primers for mouse LINE1 ORF1 were previously published41. The primers used for

DIP- and DRIP–qPCR of Alu-Y, Alu-S and LINE1s were also employed for qPCR. Gene expression was normalized by comparison to levels of GAPDH gene expression.

### 2.10.4. Validation of m6A presence on RNA:DNA hybrids using stable-isotope dilution ultra-performance liquid chromatography with tandem mass spectrometry (SID-UPLC-MS/MS).

These experiments were performed by Dr Daniel Gackowski group in Department of Clinical Biochemistry, Nicolaus Copernicus University in Toruń, Bydgoszcz, Poland. The hydrolysates were ultrafiltered prior to injection and spiked with 3 fmol/µl of internal [D3]-m6A standard in volumetric ratio 4:1. Chromatographic separation was performed using Waters Acquity UPLC system with photo- diode array detector used for quantification of unmodified nucleosides and Xevo TQ-S tandem quadrupole mass spectrometer for adenosine, m6A and [D3]-m6A. Chromatographic separation was performed using Waters CORTECS T3 (150 mm × 3 mm, 1.6 µm). The flow rate was 0.5 ml/min and the injection volume was 0.5 µl. The separation was performed with a gradient elution for 10 minutes using a mobile phase 0.05 % acetate (A) and acetonitrile (B) (0.5-3 % B for 4 min, next linearly to 70 % B in 1.5 min with subsequent column washing with 70 % acetonitrile for 1.6 min and re-equilibration with 99.5 % A for 3.6 min). All samples were analyzed in three to five technical replicates with technical means used for further calculations. The quantities of unmodified nucleosides were determined by UV detection. Mass spectrometric detection was performed using Waters Xevo TQ-S tandem quadrupole mass spectrometer equipped with an electrospray ionization source operating in positive ionization mode. Common detector parameters were as follows: capillary voltage 3.5 kV, cone voltage 15 V, collision voltage 18 V, source temperature 150 °C, nitrogen desolvation gas flow 1000 L/h, nitrogen cone gas flow 200 L/h, desolvation temperature 600 °C and nebulizer gas pressure 7 bar. Collision-induced dissociation was obtained with argon (6.0 at 3 x 10-6 bar pressure) as a collision gas. The instrument response to all compounds was

optimized by the infusion of 10 µM genuine compounds dissolved in water (10 µl/min) in mobile phase A stream, through mass spectrometer fluidics system operating in the "mixed" mode, using MassLynx 4.1 Intelli-Start feature. To provide optimum sensitivity, the most abundant fragmentation pathway for each compound was chosen (except adenosine in case of which the less abundant isotopic mass was chosen to improve the linearity of detector response): 282.2→150 (m6A), 285.2→153 ([D3]-m6A), 270.2→138 (adenosine). Stock solutions (0.05 mM) of N6-methyladenosine and stable-isotope-labelled internal standard [D3]-N6-methyladenosine (Toronto Research Chemicals) were prepared in water. Concentration of standards was measured with UV spectrophotometer and calculated by using molar extinction coefficient (ε at 266 nm = 11490). To obtain calibration curve for m6A and adenosine analyses, the ratio of area under the curve for the unlabelled nucleoside to the internal standard (A/A*) was plotted against the ratio of the molar concentrations (c/c*) of the unlabelled nucleoside to the internal standard [D3]-m6A. Absolute molar concentration of m6A was determined on the basis of isotopedilution internal calibration. The ratio of area under the curve for the unlabelled nucleoside to the internal standard (A/A*) and molar concentration of internal standard in the sample were used in calculations using calibration equations. External calibration based on peak area was used for quantitation of nucleosides using UV detection. For 2D-UPLC–MS/MS analysis of 5-methylcytidine, hydrolysates were spiked with an internal standard ([13C, D3]-5-methylcytidine, TRC Inc., Toronto, Canada) in volumetric ratio 4:1 (2 pmol/µl). Chromatographic separation was performed with a Waters Acquity 2DUPLC system and Xevo TQ-S tandem quadrupole mass spectrometer for second dimension chromatography. At-column-dilution technique was used between first and second dimension for improving retention at trap/transfer column. The columns used were: a Waters Cortex T3 column (150 mm×3 mm, 1.6 µm) at the first dimension, a Waters X-select C18 CSH (100 mm×2.1 mm, 1.7 µm) at the second dimension and Waters X-select C18 CSH (20 mm × 3 mm,

3.5 µm) as trap/transfer column. Chromatographic system operated in heart-cutting mode, indicating that selected part of effluent (2.4-2.6 min) from the first dimension was directed to trap/transfer column via 6-port valve switching, which served as "injector" for the second-dimension chromatography system. The flow rate at the first dimension was 0.5 ml/min and the injection volume was 0.5-2 µl. The separation was performed with a gradient elution for 10 minutes using a mobile phase 0.05 % acetate (A) and acetonitrile (B) (0.5-3 % B for 4 minutes, next linearly to 70 % B in 1.5 min with subsequent column washing with 70% acetonitrile and re-equilibration with 99.5 % A for 3.6 minutes). Flow rate at the second dimension was 0.3 ml/min. The separation was performed with a gradient elution for 10 minutes using a mobile phase 0.01 % acetate (A) and methanol (B) (1-50 % B for 4 minutes, isocratic flow of 50 % B for 1.5 minutes, and re-equilibration with 99 % A up to next injection). All samples were analysed in three to five technical replicates of which technical mean was used for further calculation. Mass spectrometric detection was performed using the Waters Xevo TQ-S tandem quadrupole mass spectrometer, equipped with a UniSpray ionization source with 1.2 kV potential at target pin. Collision-induced dissociation was obtained using argon 6.0 at 3 x 10-6 bar pressure as the collision gas. Transition patterns for 5-methylcytidine (258>126) and [13C, D3]-5-methylcytidine (262>130), as well as specific detector settings were determined using the MassLynx 4.1 Intelli-Start feature. To obtain calibration curves for MS/MS analyses, the ratio of the area under the curve for the unlabelled nucleoside to the internal standard (A/A*) was plotted against the ratio of the molar concentrations (c/c*) of the unlabelled nucleoside to the internal standard. Concentrations of stock solutions of standards in deionized water, used for calibration, was determined using UV-spectrometry. DNA from immunoprecipitated samples was purified using MagCore Viral Nucleic Acid Extraction Kit (cartridge code 202) according to the manufacturer's instruction. After purification, samples were completely dried in SpeedVac system and dissolved in 10 µl of MilliQ-grade deionized water. DNA was mixed

with 10 µl of buffer containing 200 mM ammonium acetate and 0.2 mM ZnCl2 (pH 4.6), nuclease P1 (0.2 U) and tetrahydrouridine (2 µg/sample) and incubated at 37° C for 1 h. Subsequently, 5.2 µl of 10 % (v/v) NH4OH and 0.26 U of alkaline phosphatase were added to each sample followed by 1h incubation at 37° C. DNA hydrolysates were spiked with a solution of internal standard in volumetric ratio 4:1, to concentration of 0.3 fmol/µl of [D3]-m6A and 200 fmol/µl of [13C, D3]- 5- methylcytidine, and analyzed as described above in 1D-mode for m6A and 2D-mode for 5- methylcytidine. Representative chromatograms and UV traces are shown below.

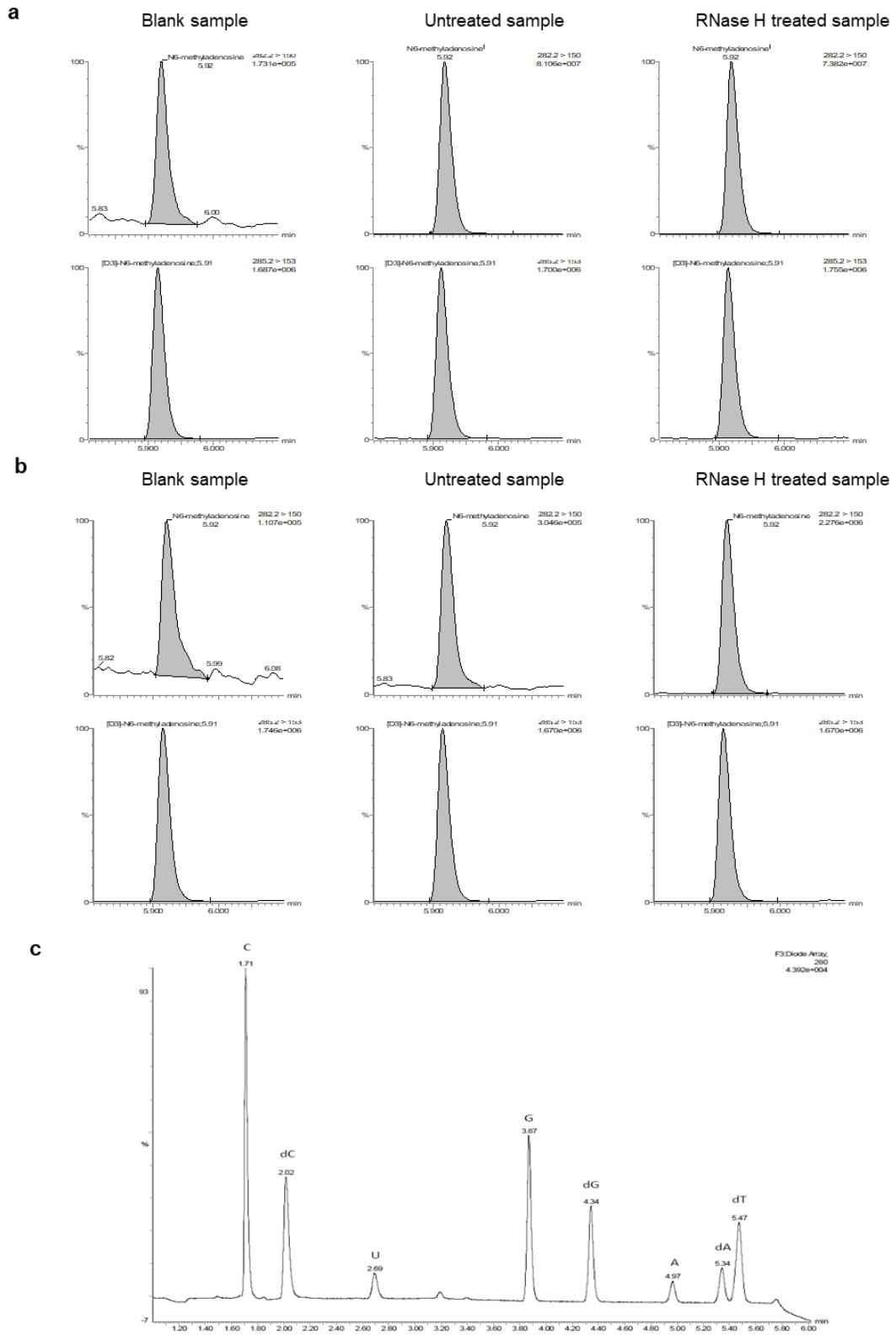



**Figure 2.10.4. SID-UPLC-MS/MS analysis of hPSCs-derived nucleic acids released and retained upon RNase H treatment.** (A) Representative chromatograms of m6A (top panel) and [D3]-m6A (bottom panel) in blank sample, untreated sample and RNase H treated sample of HUES7 hESCs-derived nucleic acids retained on 3kDa cut-off membrane. (B) Representative chromatograms of m6A (top panel) and [D3]-m6A (bottom panel) in blank sample, untreated sample and RNase H treated sample of HUES7 hESCs-derived nucleosides

passed thru 3kDa cut-off membrane. (C) Representative UV trace at 280 nm of hydrolysate of nucleic acids extracted from HUES7 cells. Similar results were obtained in 16 independent measurements with experimental samples and in 10 independent experiments during method optimisation.

### 2.11. S9.6 DRIP slot-blot

RNA:DNA hybrids and protein co-immunoprecipitation with S9.6 antibody was performed by Dr Natalia Gromak's group at Dunn's School of Pathology, Oxford University. RNA:DNA hybrids and protein co-immunoprecipitation with S9.6 antibody was performed from non-crosslinked HeLa cells as described previously (Cristini et al., 2018). Genomic DNA was isolated from HeLa cells and sonicated as previously described (Groh et al., 2014). Immunoblots of RNA:DNA hybrid immunoprecipitation samples were probed with the following antibodies: Top1 (dilution 1:2,000), YTHDF1 (dilution 1:1,000), YTHDF2 (dilution 1:500), METTL3 (dilution 1:2,000), HNRNPA2B1 (dilution 1:500) and Lamin B1 (dilution 1:2,000). Top1 and Lamin B1 were used as positive and negative controls, respectively. Images were acquired by chemiluminescence using autoradiography.

### 2.12. Chromatin Immunoprecipitation (ChIP)

For each ChIP experiments, approximately 1 x $10^7$ hPSCs grown on a matrigel-coated T75 flask in E8 medium were crosslinked with formaldehyde (1% final concentration) at room temperature. After 10 minutes of incubation, the formaldehyde was quenched with 10X Glycine followed by washing the cells twice with ice cold PBS solution containing Protease Inhibitor Cocktail II. The cells were then collected by scraping and pelleted by centrifugation at 800 x g for 5 minutes at 4 °C. Next, the cells were lysed for 15 minutes with agitation at 4 °C using cell Lysis Buffer containing Protease Inhibitor Cocktail II. Lysed cells were pelleted by centrifugation at 800 x g for 5 minutes at 4° C, discarded the debris containing supernatant and resuspended the pellets in Protease Inhibitor Cocktail II containing nuclear lysis buffer. The resulting chromatin was sonicated for 2x 15 minutes each at 4° C using Covaris S2 sonicator (Duty Cycle 5%, Intensity 4, Cycle/Burst 200). Chromatin immunoprecipitation

(ChIP) was performed using the EZ-Magna ChIP A/G Chromatin Immunoprecipitation Kit according to the manufacturer's instructions using 10 μg of anti-YTHDF2 rabbit polyclonal, 10 μg of anti-HNRNPA2B1 mouse monoclonal, or 10 μg of anti-γH2AX mouse monoclonal primary antibodies. In parallel, 5 μg of anti-Acetyl Histone H3, and normal Rabbit IgG, were run as positive and negative controls along with input. ChIP was analysed by quantitative PCR carried out with the SYBR Green PCR Master Mix (Sigma) according to standard procedures. Fold enrichment was calculated using the Pfaffl method (Pfaffl, 2001). The primers used for DRIP–qPCR were also used for ChIP–qPCR analysis.

### 2.13. Electrophoretic Mobility-Shift Assay (EMSA)

EMSA technique for assessing *in vitro* binding of YTHDF2 to various m6A-modified and nonmodified oligonucleotides was performed.

#### 2.13.1. Purification of recombinant YTHDF2.

Full-length YTHDF2 was cloned into pET-28b with N-terminal His-tag (Genescript). The plasmid was expressed in BL21(DE3) cells and incubated with LB-medium (Puls Medical, 244610) with sorbitolat at 37 °C until OD600=0.7. Expression was induced with 300 μM of isopropyl-b-d-thiogalactoside overnight at 18 °C. Cells were pelleted by centrifugation at 3,000g for 10 min at 4 °C, resuspended in lysis buffer (50mM Tris pH 8.0, 300mM NaCl, 10mM β-mercaptoethanol (ME), 10 mM imidazole) and sonicated. After sonication, the extract was centrifuged at 19,000 g for 20 min at 4 °C. The supernatant was loaded to Protino Ni-NTA agarose prepared as described by the producer (Macherey-Nagel, 745400.100) in a 50-ml tube and incubated at 4 °C with rotation for 30 min. After centrifugation at 3,000g for 2 min, the Ni-agarose-bound YTHDF2 was washed with buffer containing 50mM Tris pH 8.0, 300 mM NaCl, 10 mM ME and 50 mM imidazole. Recombinant YTHDF2 was eluted with five washes

(1.5 ml each) of elution buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM ME and 300 mM imidazole).

### 2.13.2. Formation of end-labelled synthetic RNA:DNA hybrids

To generate RNA:DNA hybrids, the RNA and DNA oligonucleotides were mixed in equimolar concentrations, incubated for 20 min at 98º C, slowly cooled down in a heating block and placed on ice. Oligonucleotides were end-labelled with γ32P-ATP using T4 polynucleotide kinase (NEB).

### 2.13.3. Electrophoretic Mobility-Shift Assay (EMSA).

Approximately 10-300 ng of His-fused human recombinant YTHDF2 or 30-300 ng of human recombinant FLAG-fused YTHDF2 (Active Motif, catalogue number 31573) were incubated with 0.15 pmol of the labelled RNA:DNA hybrid or single stranded RNA probes in 15 µl of binding buffer containing 1 µg poly(dI-dC) (Sigma) for 30 min on ice. Either high salt (80-150 mM NaCl) binding buffer (12.5 mM Hepes pH 7.9, 5 mM MgCl2, 0.5 mM EDTA, 80/150 mM NaCl, 1 mM DTT, 0.025 % NP-40, 50 µg/ml BSA, 15 % glycerol) or low salt binding buffer previously used for methyl-CpG-binding proteins and DNA substrates24 (6 mM Hepes pH 7.9, 4 mM Tris HCl pH 8.0, 30 mM KCl, 1 mM EDTA, 1 mM DTT, 30 µg/ml BSA, 5% glycerol) were employed for EMSA. The binding reactions were separated by electrophoresis in 5 % acrylamide gel and 0.5 × TBE buffer at 4º C. The signal was visualised using phosphoimager.

### 2.14. Microscale thermophoresis.

Microscale thermophoresis was employed to study the interaction of YTHDF2 with modified or unmodified RNA and the RNA:DNA hybrid synthetic substrates used in EMSA experiments. Purified fulllength YTHDF2 with N-terminal His-tag was labelled with the NT-647 RED-trisNTA dye for His-tagged proteins following the recommended labelling protocol (Nanotemper Technologies). One hundred microliters of 200-nM protein in PBS buffer was

mixed with 100 µl of 100 nM dye in the supplemented PBS-T buffer and incubated at room temperature for 30 min. The labelled protein was then added to a 1:1 dilution series of the respective substrate in UltraPure DNase/ RNase-Free Distilled Water (Thermo Fisher Scientific) to a final concentration of 50nM. The different protein-substrate samples were loaded into standard NT.115 MST capillaries (Nanotemper). Fluorescence profiles were measured at 25 °C in a Monolith NT.115 instrument using the red channel. Data were collected at either 20% (single-strand substrates) or 40% (hybrid substrates) MST power and 60% excitation power. Changes in fluorescence (ΔFnorm) due to thermophoresis were measured as the signal difference between time points 0 and 10 s. Data were normalized and plotted as a function of ligand concentration, and a binding curve was fitted to the average of six independent dilution series of each substrate.

**2.15. Statistics:**

At least two, and typically three, independent experiments were carried out for all experiments unless indicated otherwise. DRIP and DIP were performed in two, and RNA-seq in three, biologically independent experiments. All experiments were replicated independently. Statistical tests used for individual experiments are described in the corresponding figure legends. For quantification of m6 A, γH2AX and S9.6 signal intensities, statistical significance was determined using either unpaired two-tailed Student's t-test or unpaired two-tailed Welch′s t-test. Signal intensity and qPCR data were plotted and analyzed in GraphPad Prism v.7.04.

# Chapter 3

# Discovery of N6-methyladenosine on the RNA components of RNA:DNA hybrids (R-loops) in mammalian cells.

## Chapter 3: Results

### 3. Discovery of N6-methyladenosine on the RNA components of RNA:DNA hybrids (R-loops) in mammalian cells.

N6-methyldeoxyadenosine (6mA) is a pervasive DNA modification whose primary function in prokaryotes is thought to permit host genomes distinguish self from invading DNA as part of the restriction-modification system (Dunn and Smith, 1955; Vanyushin et al., 1970; Iyer et al., 2011). Comparative genomic analysis of DNA modifying enzymes postulated the lateral transfer of adenine methylases to eukaryotic systems (Iyer et al., 2011). Although evidence of this modification occurring in simple eukaryotes such as *Tetrahymena* was known since the 1970s (Gorovsky et al., 1973), it is the more recent reports of 6mA presence in multi-cellular eukaryotes (Greer et al., 2015a; Zhang et al., 2015; Fu et al., 2015) that re-ignited investigations into the biological functions of 6mA beyond the restriction-modification system in bacteria. However, contradictory reports of 6mA detection in mammalian genomes (Wu et al., 2016a; Schiffers et al., 2017) raised new doubts regarding the presence of 6mA in vertebrates (Douvlataniotis et al., 2020). In addition to this, both bacterial and mammalian transcriptomes contain abundant yet dynamic methylated adenine (m6A) on RNA (Jia et al., 2011; Meyer et al., 2012; Dominissini et al., 2012; Deng et al., 2015). Moreover, the presence of high levels of 6mA in *Mycoplasma* (Lluch-Senar et al., 2013) often found in mammalian cells cultured in non-sterile conditions further complicates interpretation of studies reporting the presence of 6mA in mammalian genomes (Douvlataniotis et al., 2020). Therefore, careful re-evaluation of reported 6mA in mammalian genomes is warranted. As the amount of 6mA reported in eukaryotes is several folds lower than in prokaryotes (O'Brown and Greer, 2016), it was reasoned that if 6mA is present in mammalian DNA, its content is significantly lower than the levels of the well-studied DNA 5mC. Herein, a highly sensitive immunochemical technique, optimized, specifically for the analysis of low-abundance DNA modifications

(Abakir et al., 2016) was employed with an anti-6mA specific antibody (figure 3.1A) to investigate the presence and genomic distribution of 6mA in human pluripotent stem cells (hPSCs).

### 3.1. m6A is detectable by immunochemistry and LC-MS/MS in hPSCs RNA

Before investigation of 6mA presence in hPSCs genomes, first, we examined the presence of RNA m6A in hPSCs using a highly sensitive immunochemical technique, optimized, specifically for the analysis of low-abundance nucleic acid modifications (Figure 3.1A) (Abakir et al., 2016) with anti-m6A specific antibody in non-denaturing condition. Importantly, under these conditions, genomic 6mA cannot be detected. This is due to the omission of DNA denaturation (e.g. using 4M HCl) step that is essential and routinely used for immunostaining of DNA modifications (Santos and Dean, 2006; Wheldon et al., 2014; Mondo et al., 2017). Immunostaining of REBL-PAT human pluripotent stem cells (hPSCs) with this antibody showed prominent nuclear and cytoplasmic m6A signal (Figure 3.1B, top panel) that disappeared in RNase A pre-treated control hPSCs (Figure 3.1B inset panel). These results confirmed the presence of abundant m6A in human pluripotent stem cells (Dominissini et al., 2012) and importantly, validated the suitability of the immunostaining technique for m6A detection in human cells. Confirming an RNA origin of m6A staining signal in human cells, m6A immunostaining of hPSCs in presence of DNase I, an endonuclease that degrades DNA (Salnikow et al., 1970) led to a loss of DAPI signal but preserved the characteristic m6A staining pattern in untreated control cells (Figure 3.1C). In agreement with these results, using liquid chromatography-tandem mass spectrometry (LC-MS/MS), we showed that approximately 0.2% of total adenosine residues isolated from hPSCs were m6A modified (Figure 3.1D).

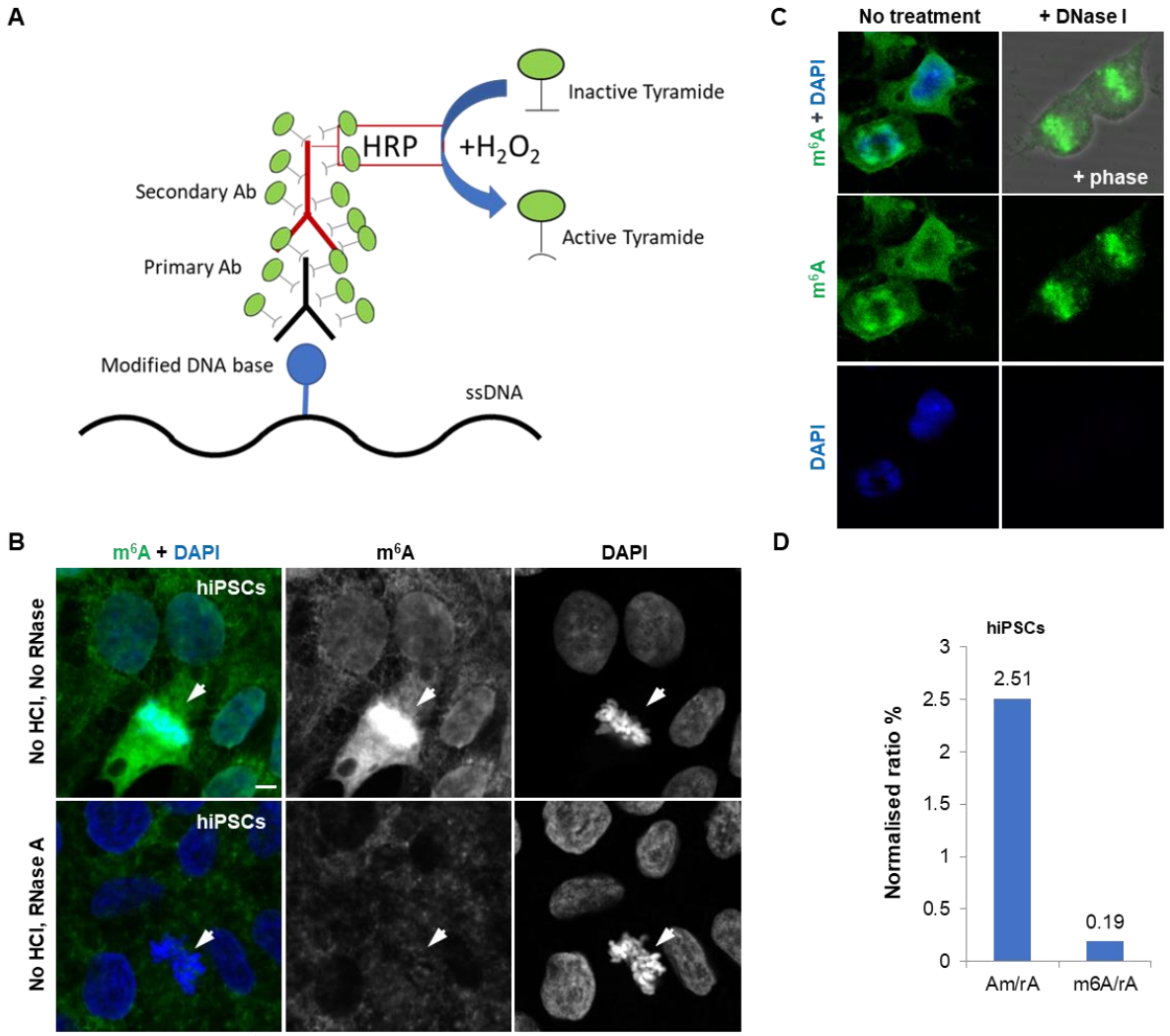



**Figure**. **3.1**. **Detection of RNA m6A in hPSCs using immunochemistry and LC-MS/MS.** (A) a schematic of the Tyramide Signal Amplification for immunofluorescent enhancement employed here for the detection of low-abundance non-canonical nucleic acid bases, adapted from (Faget and Hnasko, 2015). (B) REBL-PAT hPSCs co-immunostained for m6A (green) and DAPI (blue) in non-denaturing conditions, either pre-treated with RNase A (Top panel) or untreated (bottom panel). Merged images and individual channels are shown. Dividing cells are arrowed (C) Co-immunostaining of m6A and DAPI in REBL-PAT hPSCs performed in non-denaturing conditions with or without DNase I treatment as indicated. Merged images consisting of m6A (green) and DAPI (blue) channels together with phase in case of DNase I treated sample) and individual channels are shown. Scale bars are 10 μm. (D) The ratios of the indicated ribonucleotides obtained from the quantification of LC-MS/MS peaks in the total hPSCs RNA.

### 3.2. Detection of RNase A-resistant m6A immunostaining signal in mitotic chromatin of various human cell lines.

To examine whether 6mA is detectable in human cells, hPSCs were co-immunostained with m6A and 5mC following pre-treatment with 4N HCl. Whereas 5mC immunostaining signal was restricted to hPSCs nuclei, m6A signal was observed throughout stained cells but most prominent in the nuclei (Figure 3.2A). In order to discriminate the ubiquitous m6A signal from the potential presence of low-abundance 6mA, a parallel m6A co-immunostaining with 5mC was performed in mitotic cell-enriched hPSCs pre-treated with RNase A. Remarkably, a strong immunostaining signal in mitotic chromatin potentially originated from 6mA was detected in RNase A treated hPSCs (Figure 3.2B). As conclusive detection of 6mA in mammalian cells remain elusive (Douvlataniotis et al., 2020; Schiffers et al., 2017), detection of 6mA in hPSCs would have several implications on studies of chromatin biology, cellular differentiation and development. Therefore, the same immunostaining was repeated in various cancer cell lines including LN18 (Figure 3.2C) and HeLa cells (Figure 3.2D -E). In both cell lines, a very similar RNase A-resistant mitotic m6A staining signal was observed. These results seemingly indicated the presence of 6mA in the chromatin of human cells. In an attempt to verify these results, next, genomic DNA isolated from both mitotic cells synchronised using colcemid (Rieder and Palazzo, 1992) or asynchronous hPSCs lines (REBL-PAT and HUES7) was examined for the presence of 6mA using liquid chromatography–tandem mass spectrometry (LC–MS/MS). In parallel, other modified DNA bases known to decorate genomes of hPSCs (5mC; 5-hydroxymethylcytosine, 5hmC and 5-formylcytosine, 5fC) were also measured in the same samples. Astonishingly, LC-MS/MS analysis showed no 6mA in these cells even at levels comparable to that of 5fC [about 20 parts per million cytosine, (Pfaffeneder et al., 2011)] (Figure 3.2. F). Although these results are consistent with a growing list of studies that failed to report appreciable 6mA detection in mammalian genomes (Douvlataniotis et al., 2020;

Schiffers et al., 2017), nevertheless, these results highlighted a massive discrepancy with the immunostaining results obtained previously and shown in (Figure 3.1B).

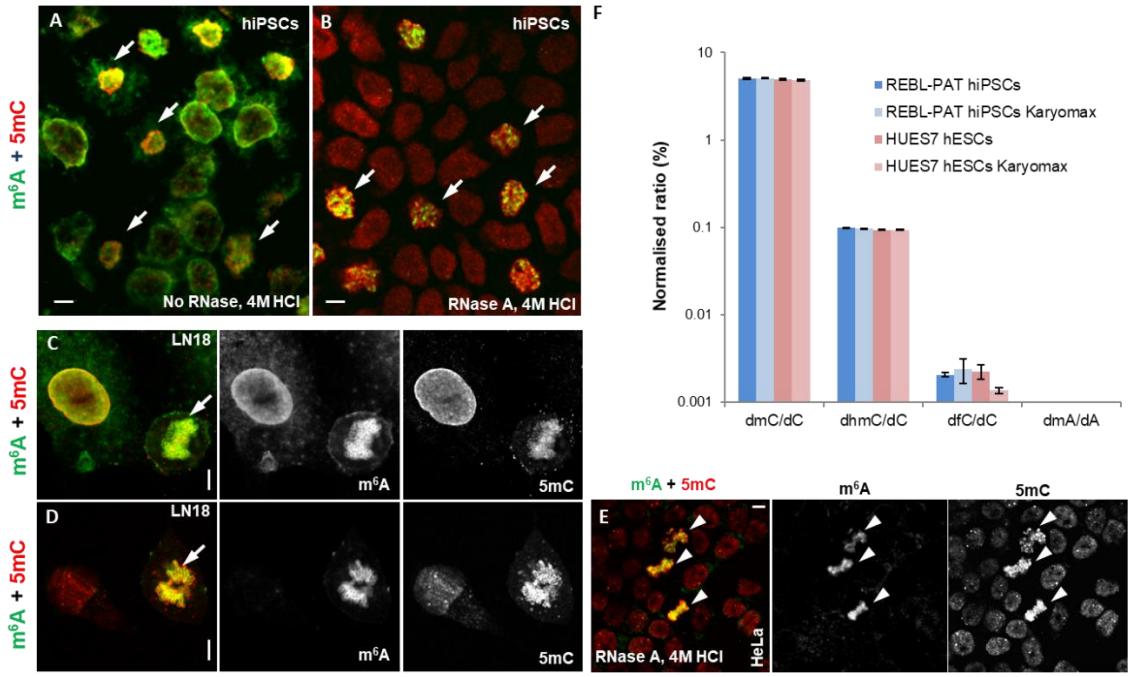



**Figure 3.2. Mitotic chromatin of different human cell lines exhibit RNase A-resistant m6A immunostaining signal.** m6A and 5mC co-immunostaining of non-RNase A treated (A) and RNase A treated (B) hPSCs mitotically enriched using colcemid treatment with addition of 4M HCl-denaturing step. Representative co-immunostaining of m6A and 5mC in non-RNase A treated (C) and treated (D) LN18 cancer cells pre-treated with 4M HCl-denaturing step. Co-immunostaining of m6A and 5-methyldeoxycytosine (5mC) in RNase A treated HeLa cells with the addition of 4M HCl-denaturation step (E). Merged images are shown (A-B). Merged views and individual channels are presented in (C, D and E). Mitotic cells are indicated with arrows (A-E). (F) Ratios of the indicated deoxynucleotides obtained from the quantification of LC–MS/MS peaks in KaryoMAX-treated and untreated hiPSC/hESC DNA. Data are means ± s.d., n = 2 MS experiments.

### 3.3. Validating the specificity of anti-m6A antibody with synthetic 6mdATP in hPSCs.

The striking discrepancy between the strong m6A immunostaining results and the lack of 6mA in LC-MS/MS analysed hPSCs genomes in (Figure 3.2), prompted the re-evaluation of anti-6mA antibody specificity using immune-competition experiments. As previously observed, strong m6A staining was observed in the absence of any competing nucleotide (Figure 3.3A). Similarly, neither the presence of 100 nM deoxyadenosine triphosphate, (dATP) (Figure 3.3B), nor 100 nM deoxythymidine triphosphate, dTTP (Figure 3.3C), interfered with the m6A staining signal. In contrast, co-immunostaining of hPSCs for m6A and 5hmC in the presence of only 10 nM N6-methyldeoxyadenosine triphosphate, 6mdATP, resulted in the specific loss of m6A signal without interference with 5hmC signal (Figure 3.3D). Though these experiments confirmed the specificity of the anti-m6A antibody (Mondo et al., 2017) in the immunostaining experiments above, to this point, the source of the m6A immunostaining signal in mitotic chromatin of human cells remained a mystery.

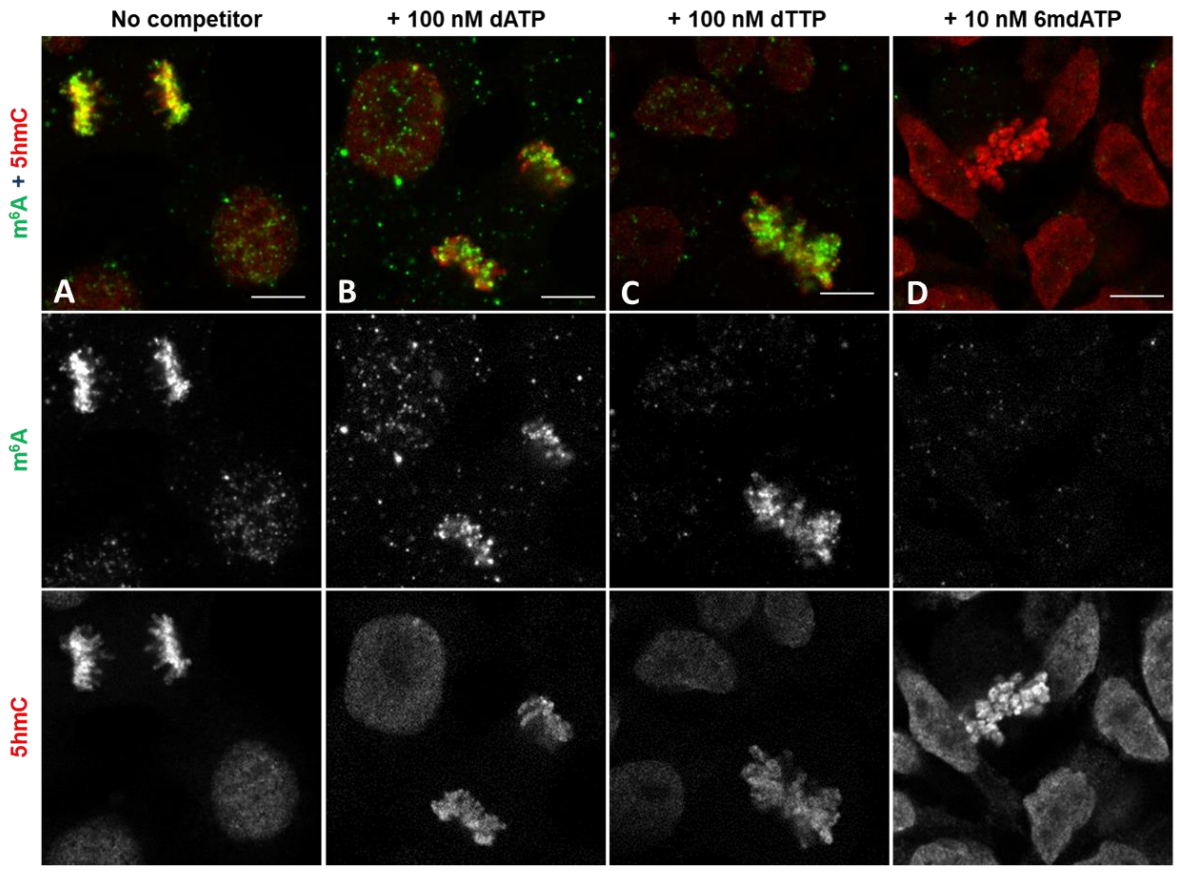


**Figure 3.3. m6A immunostaining in hPSCs is specifically competed by 6mdATP but not by unmodified deoxythymidine triphosphate**. Representative confocal microscopy images of RNase A treated hPSCs co-immunostained for m6A and 5-hydroxymethylcytosine (5hmC) following 4M HCl treatment. The staining was performed either without competitor (A), or in the presence of deoxyadenosine triphosphate (dATP)(B), deoxythymidine triphosphate, dTTP (C) and N6-methyldeoxyadenosine triphosphate, 6mdATP(D). Scale bars are 20 μm.

### 3.4. Discovery of m6A presence on the RNA component of RNA:DNA hybrids (R-loops)

If the anti-m6A antibody is specific for m6A as shown in (Figure 3.3), then how is it possible that 6mA was not detectable with the highly sensitive LC-MS/MS (Figure 3.2)? This led to the hypothesis that the prominent m6A immunostaining signal in RNase A treated hPSCs might originate from an m6A bearing RNA trapped in a hybrid with a DNA as part of the R-loop. R-loops are three-stranded nucleic acid structures composed of an RNA:DNA hybrid and a displaced single stranded DNA (Thomas et al., 1976; Drolet et al., 1995; Santos-Pereira and Aguilera, 2015). Despite their unusual structure, R-loops play crucial functions in several physiological process that include DNA replication (Masukata et al., 1987), Ig class switch recombination in stimulated B-cells (Yu et al., 2003), countering DNA methylation (Grunseich et al., 2018) and, as shown more recently, transcriptional regulation (Ginno et al., 2012; Skourti-Stathaki and Proudfoot, 2014; Niehrs and Luke, 2020). Despite their widely recognised roles in cellular physiological processes, unless resolved, formation of unscheduled R-loops compromise genome integrity (Aguilera and García-Muse, 2012; Reijns et al., 2012; Crossley et al., 2019). Therefore, organisms from all three domains of life utilise an endoribonuclease, RNase H, to resolve RNA:DNA hybrids by specifically degrading the RNA component of the hybrids (Drolet et al., 1995; Gaidamakov et al., 2005). Consequently, sensitivity to RNase H has become one of the defining features of RNA:DNA hybrids that is routinely used to investigate the presence of these structures (Drolet et al., 1995; El Hage et al., 2014; Sanz et al., 2016b). In addition, the generation of anti-RNA:DNA hybrid monoclonal (S9.6) antibody (Boguslawski et al., 1986) proved crucial in isolation and mapping the genome-wide distribution of RNA:DNA hybrids (Ginno et al., 2012; El Hage et al., 2014). To test whether

the mitotic chromatin-associated m6A immunostaining signal originate from the RNA component of R-loops, hPSCs were co-immunostained with m6A and the mitosis marker, phosphorylated Histone H3 (Ser10) either without RNase treatment (Figure 3.4A), pre-treated with RNase A only (Figure 3.4B) or following pre-treatment with RNase H (Figure 3.4C). As expected, mitotic chromatin-associated m6A signal was detected in the absence and presence of RNase A (Figures 3.4A-B). Remarkably, mitotic chromatin m6A immunostaining signal completely disappeared or significantly decreased in RNase H treated hPSCs (Figure 3.4C). The loss of m6A immunostaining signal in RNase H treated hPSCs provided the first ever evidence of a non-canonical nucleic acid base in R-loops. To verify these results using an independent approach, genomic DNA of two hPSCs lines (REBL-PAT human induced pluripotent stem cells and HUES7 human embryonic stem cells) were isolated either in the presence of RNase H or without it, before the analysis of m6A levels in both factions using stable-isotope dilution ultra-performance liquid chromatography with tandem mass spectrometry (SID-UPLC-MS/MS). Whereas no significant change was observed in the levels of ribo-5-methylcytidine, ribo-$m^5C$, following RNase H treatment, we observed more than fivefold increase in the amount of m6A recovered from the RNase H pre-treated fraction compared to m6A levels in the untreated genomic DNA of hPSCs.

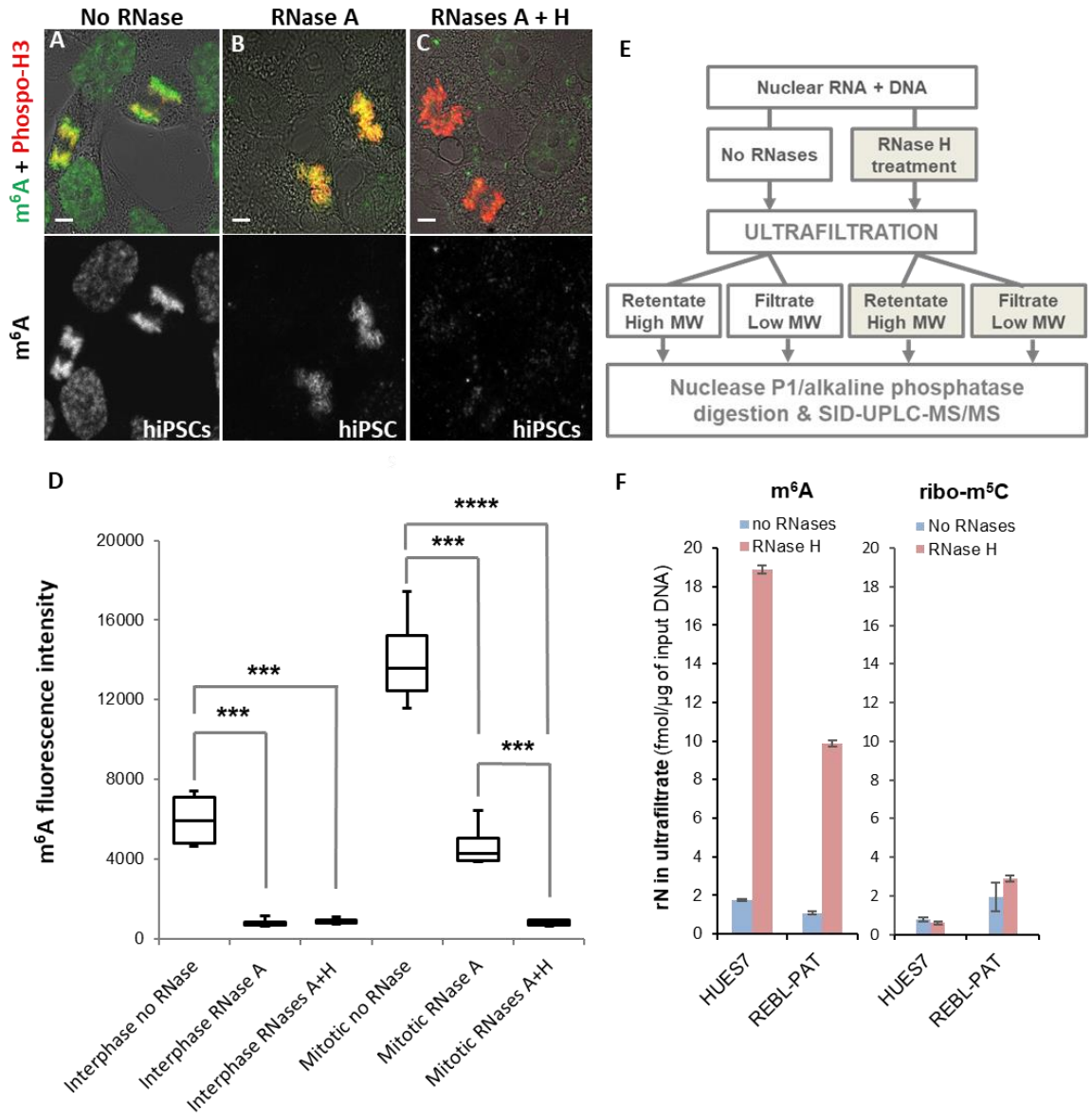



**Figure 3.4. m6A marks the RNA component of RNA:DNA hybrids in hPSCs.** Representative confocal microscopy images of hPSCs co-immunostained for m6A and phospho-Histone H3 (phosphorylated histone 3 at serine 10) without RNases (A) and following RNase A only treatment (B) or combined RNases A + H treatment. Merged views and m6A channel are presented. Scale bars, 5µm. (D) Box plots showing quantification of m6A signal intensity in representative 20 cells each, interphase and mitotic following no RNase treatment, RNase A only treatment and combined RNase A and H treatment. Significance was determined using unpaired two-tailed Student's t-test (***p<0.0001). (E) Schematic illustrating design of the experiment on SID–UPLC–MS/MS analysis of hPSCs-derived genomic DNA released (filtrate) and retained (retentate) following RNase H treatment. (F) SID–UPLC–MS/MS quantification of m6A and ribo-m5C in the fractions of HUES7 hESC- and REBL-PAT hiPSC-derived genomic DNA released following RNase H treatment. Data are shown as means ± s.d., n = 4/n = 3 MS experiments for m6A/ribo-m5C quantification. rN, ribonucleosides.

### 3.5. Summary and conclusion.

Taken together, these results constitute the first report of m6A presence on the RNA component of RNA:DNA hybrids. However, the genomic distribution of m6A marked RNA:DNA hybrids was not known, insights on which can provide clues regarding the biological significance of marking R-loops with m6A. Given the dynamic nature of co-transcriptionally deposited m6A on nascent RNAs (Hsu et al., 2017; Ke et al., 2017), can the function of m6A on R-loops be uncovered by modulating its incorporation and removal?. The remainder of the thesis aimed to answer these questions.

# Chapter 4

# Genome-wide analysis reveals that majority of RNA:DNA hybrids in hPSCs are marked by m6A

## Chapter 4: Results

### 4. Genome-wide analysis reveals that majority of RNA:DNA hybrids in hPSCs are marked by m6A.

### 4.1. Development of a new protocol for genome-wide mapping of m6A-modified RNA:DNA hybrids

Following the generation of S9.6 (a monoclonal anti-RNA:DNA hybrid) antibody (Boguslawski et al., 1986), immunoprecipitation of RNA:DNA hybrid has widely been used to examine genomic sequences containing these structures (El Hage et al., 2010; Groh et al., 2014). By incorporating next generation whole-genome sequencing technologies, genome-wide maps of RNA:DNA hybrids (S9.6 DNA/RNA precipitation, DRIP) has been profiled in several species, ranging from yeast to human (Ginno et al., 2012; Nadel et al., 2015; Wahba et al., 2016; Dumelie and Jaffrey, 2017). However, as the presence of m6A on RNA:DNA hybrids was not previously known, understandably, there was no m6A DRIP protocol available. Therefore, in this study, the previously published S9.6 DRIP-seq approach (Ginno et al., 2012; Chan et al., 2014) was specifically adapted for mapping m6A modified RNA:DNA hybrids; whereby the anti-m6A antibody (Mondo et al., 2017) replaced the S9.6 antibody with the additional step of heat denaturation (to permit m6A detection) and identification of m6A and S9.6 peaks by comparing to RNase H treated controls (Figure 2.7.8).

### 4.2. Validation of the m6A DRIP approach

As the m6A DRIP protocol (Figure 2.7.8) has previously not been validated, first, the specificity and affinity of S9.6 antibody for m6A-modified RNA:DNA hybrids was examined using m6A modified and non-modified 81-mer synthetic single stranded RNA oligonucleotides hybridized with complementary DNA oligonucleotides (Figure 4.2A). Interestingly, evaluation of the immunoprecipitation results by qPCR revealed comparable efficiency of S9.6 antibody towards m6A modified and non-modified RNA:DNA hybrids, indicating that the presence of m6A on the hybrids did not hinder their immunoprecipitation by the S9.6 antibody (Figure

4.2B). Performing the same qPCR evaluation by replacing the S9.6 antibody with anti-m6A antibody showed the specific enrichment of m6A-modified RNA:DNA hybrids in m6A DIP analysis (Figure 4.2C). In contrast, m6A DIP-qPCR analysis did not reveal any enrichment of non-m6A containing RNA:DNA hybrids, single stranded RNA with or without m6A and single stranded DNA (Figure 4.2C). These results revealed the specificity of the m6A antibody in DRIP approaches; thus, confirming the validity of the adapted m6A DRIP approach for genome-wide mapping of m6A-marked R-loops. Intriguingly, when compared to S9.6, the anti-m6A antibody had appreciably greater pull-down efficiency of the same synthetic hybrid substrate, suggesting the two antibodies have markedly different affinities to the same substrate (Figure 4.2D).

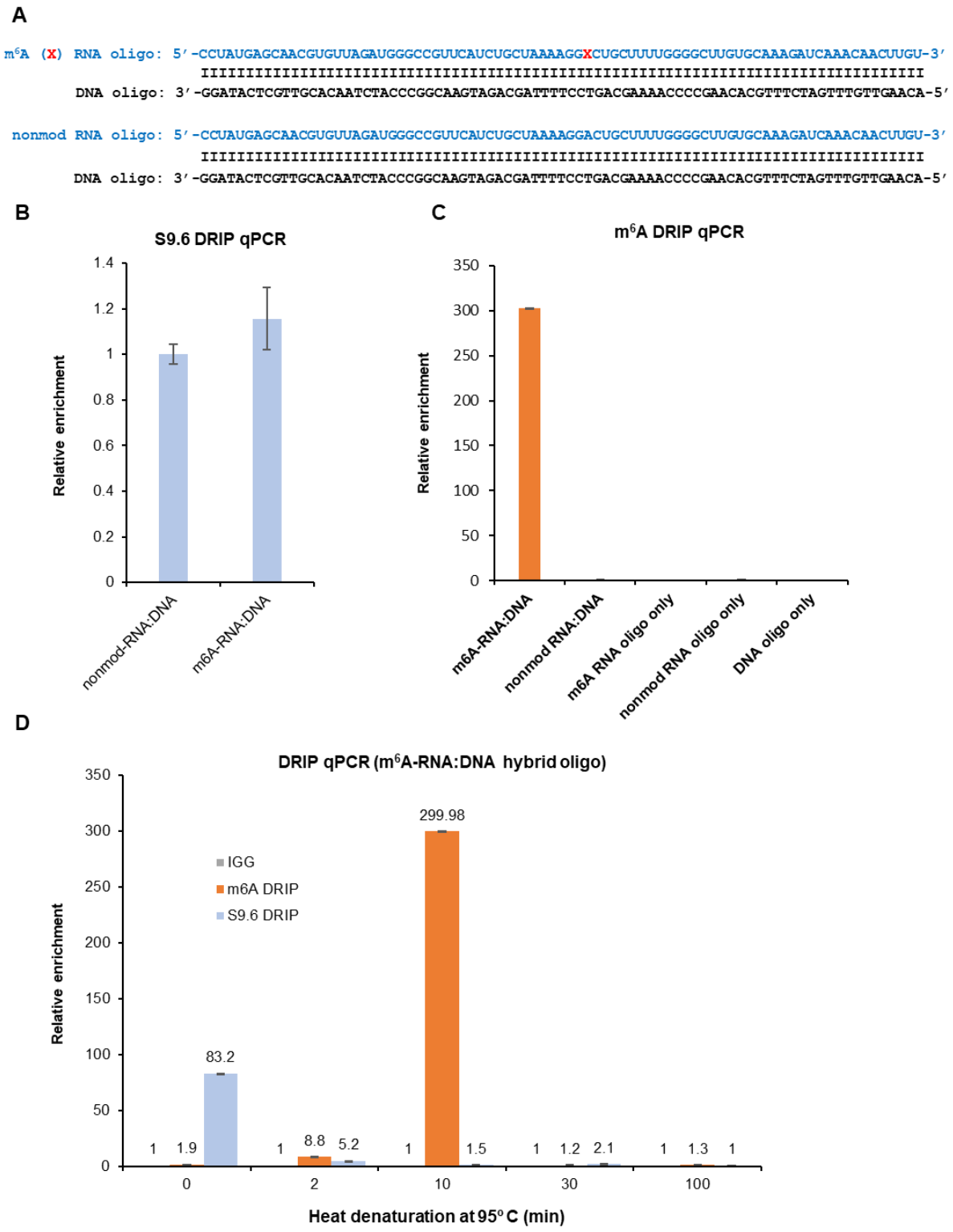



**Figure 4.2. Validation of m$^6$A DIP using synthetic RNA and DNA substrates** (A) Schematic demonstrating the sequences of synthetic oligonucleotides and RNA:DNA hybrids used for spike in validation experiments. (B) S9.6 DRIP exhibits comparable efficiencies in precipitating non-modified and m$^6$A-containing RNA:DNA hybrids. Relative enrichment of DRIP performed on the m$^6$A-containing RNA:DNA synthetic substrate normalized against that of DRIP done on equivalent amount (0.1 pmol) of non-modified spike in synthetic

RNA:DNA hybrid. (C) The results of $m^6A$ DIP on 0.1 pmol of the indicated spike in synthetic oligonucleotides and RNA:DNA hybrids. Unlike non-modified RNA:DNA hybrid substrate or single stranded $m^6A$-containing RNA oligonucleotide, $m^6A$-containing RNA:DNA hybrid is efficiently detected by $m^6A$ DIP technique. (D) The relative enrichment of DIP/DRIP performed on 0.1 pmol of spike in synthetic $m^6A$-containing RNA:DNA hybrid (81-mer with 46 % GC content) shown in (A) using anti-$m^6A$ and S9.6 antibodies along with IGG for IP after the indicated time intervals of heat denaturation. Data are means ± SD, n = 3 independent experiments.

### 4.3. Genomic distribution of m6A modified RNA:DNA hybrids

To map the genome-wide distribution of m6A marked R-loops, m6A and S9.6 DRIP reactions were performed in parallel on hPSCs-derived genomic DNA followed by high-throughput sequencing of the DNA template as described in (Figure 4.1). As controls, m6A and S9.6 DRIP reaction were performed in parallel on hPSCs genomic DNA pre-treated with recombinant *E. coli* RNase H (Figure 4.1). In collaboration with the ADAC team at the University of Nottingham, the DRIP-seq reads were mapped to the human genome, followed by peak calling and identification of the consensus peaks after analysis of the replicate correlation of the identified peaks for the obtained datasets (4.3.1A-B). Amongst the consensus peaks identified, there were 23,582 m6A and 6,713 S9.6 peaks respectively, of which 6,512 were overlapping peaks (Figure 4.3.2A). Importantly, both m6A and S9.6 peaks displayed almost identical distribution patterns across various genomic regions (Figure 4.3.2B) and the majority of these peaks were RNase H sensitive (Figure 4.3.2C), indicating the presence of m6A on the RNA component of RNA:DNA hybrids.

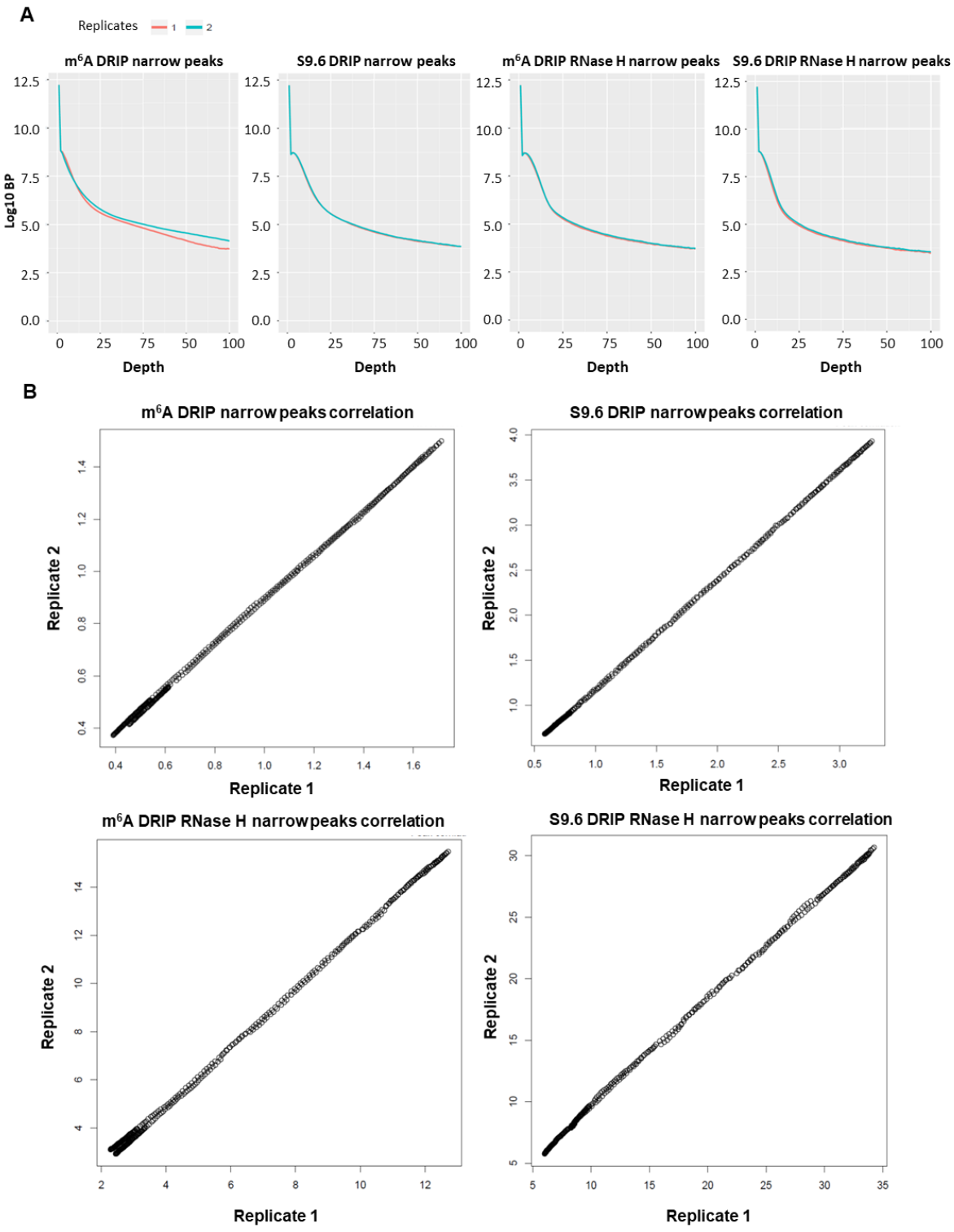



**Figure 4.3.1. Replicate correlation analysis for m6A and S9.6 DRIP-seq peaks.** (A) Replicates' coverage histograms for the narrow peaks of the indicated DRIP/DIP samples. (B) Peaks correlation histograms for the indicated samples.

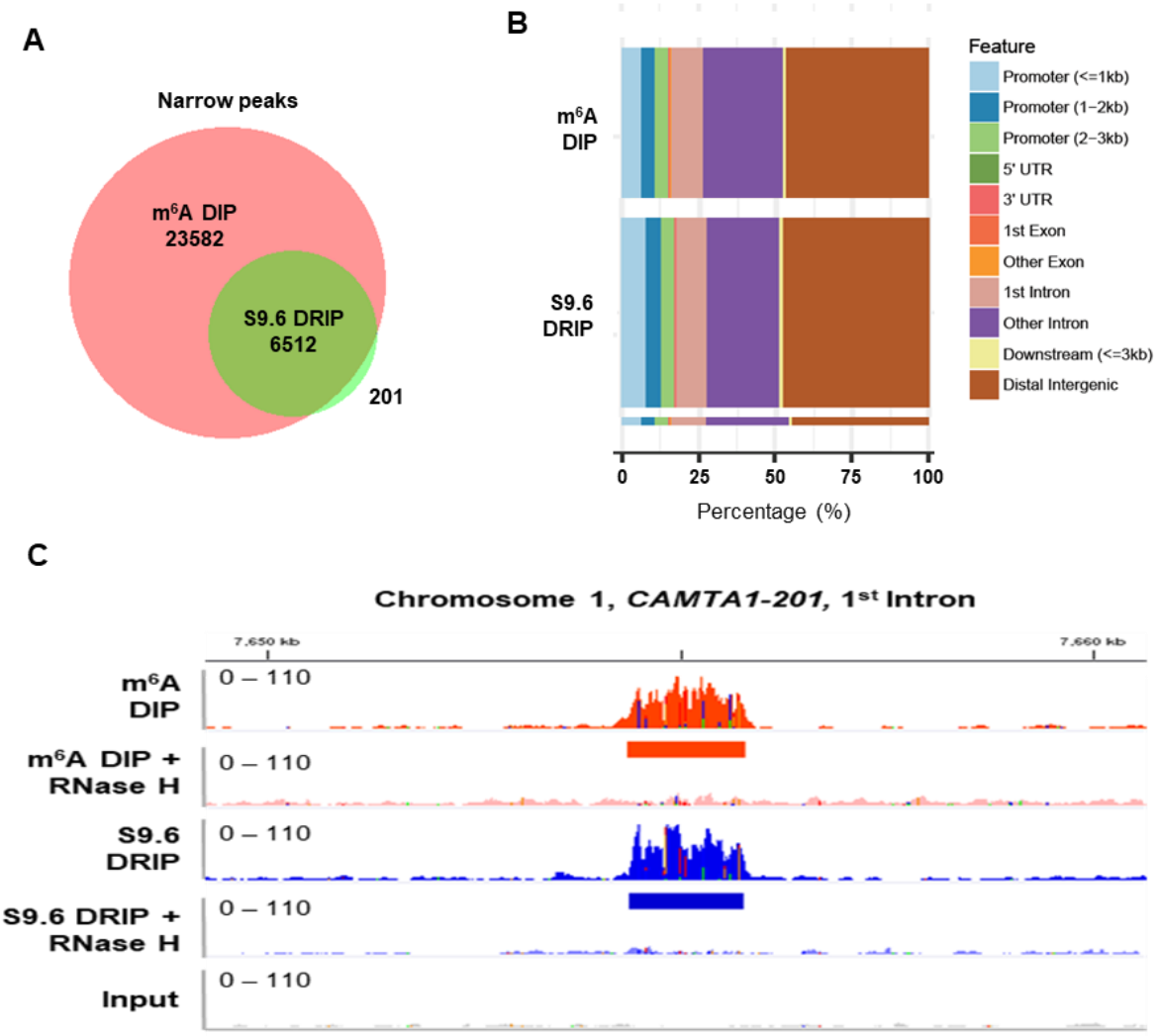



**Figure 4.3.2. m6A is present on the majority of RNA:DNA hybrids in hPSCs**. (A) Venn diagram illustrating overlap between m6A DIP and S9.6 DRIP consensus peaks in REBL-PAT hiPSCs (B) Distribution of m6A and S9.6 peaks at the indicated genomic features in hiPSCs. (C) Coverage plots of m6A DIP and S9.6 DRIP densities (counts per kilobase, CPK) in the intron of the CAMTA1-201 gene. m6A DIP and S9.6 DRIP peaks are marked by red and blue rectangles, respectively.

### 4.4. Analysis of m6A and S9.6 peak distribution across various genomic features.

Consistent with co-transcriptional generation of R-loops (Stork *et al*., 2016), m6A peaks across different gene categories accumulated mostly in protein coding regions, antisense transcripts and long noncoding RNAs (Figure 4.4A). Regulatory regions, CTCF-binding sites, transcription factor (TF) binding sites, open chromatin, and enhancer sequences showed enrichment of m6A peaks (Figure 4.4B). Similarly, highest levels of RNase H sensitive peaks

were detected in introns, intergenic regions, 3 kilobases (Kb) up- and downstream of coding sequences (CDS) followed by less peaks in promoters, exons and promoter flanking regions (Figure 4.4C-D). In repetitive sequences, both m6A and S9.6 peaks showed enrichment in Long Terminal Repeat-containing (LTR) retrotransposons, and non-LTR retrotransposons such as LINE1 and Alu sequences but neither peaks were detected in α-satellite repeats (Figure 4.4E). These results confirm the presence of m6A on the majority of RNA:DNA hybrids.

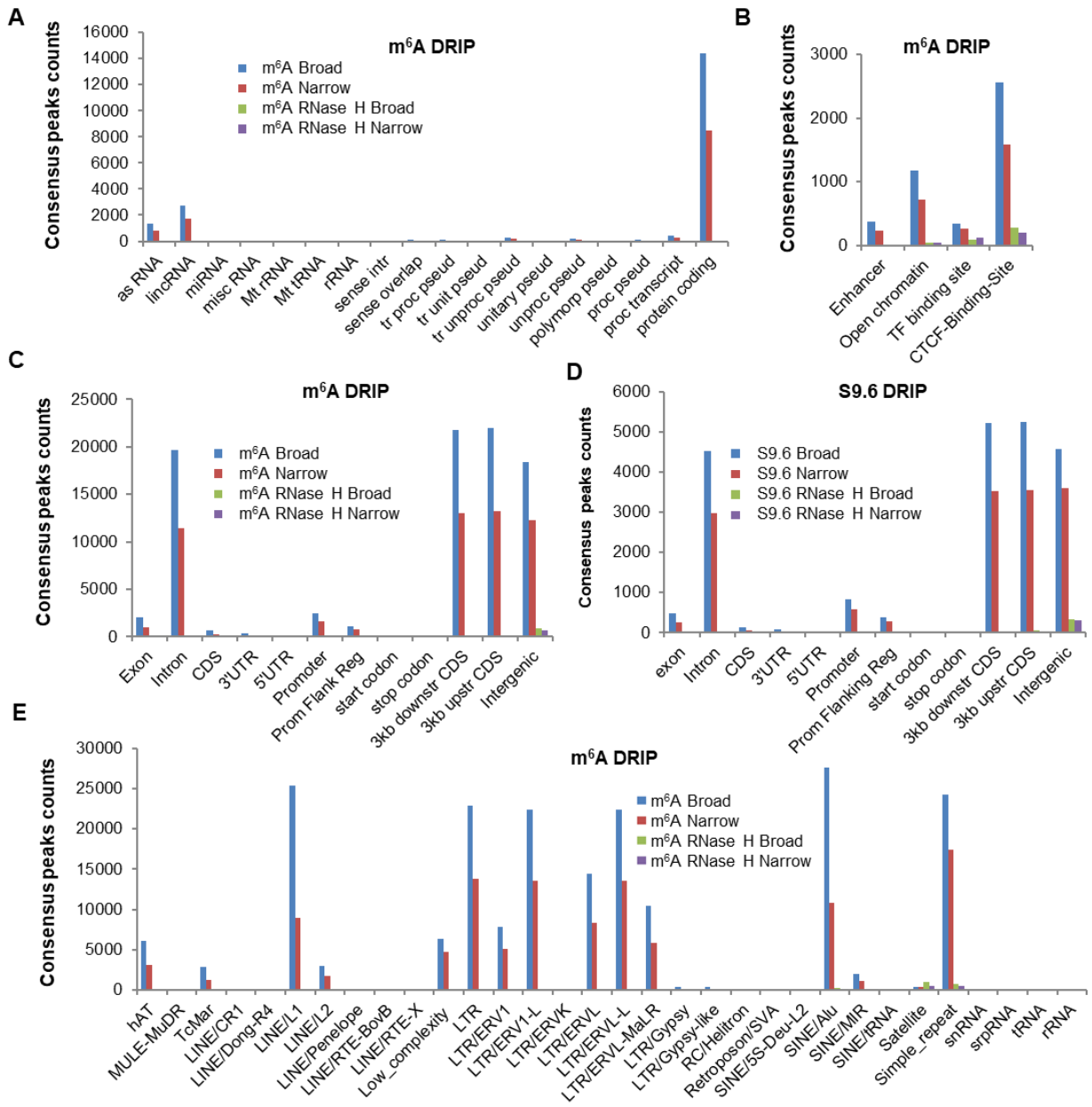



**Figure 4.4. Analysis of m6A and S9.6 peak distribution across distinct genomic features.** (A, B) Distribution of the m$^6$A DRIP consensus peaks counts across the indicated gene

categories (A) and regulatory sequences (B). (C-D) Distribution of the $m^6A$ (A) and S9.6 DRIP (B) consensus peaks counts across the indicated genomic features. (E) Distribution of $m^6A$ DRIP consensus peaks counts at the indicated repetitive sequences.

### 4.5. Summary and conclusion

Since the genome-wide distribution of m6A DRIP-seq datasets exhibited almost identical patterns to that of S9.6 DRIP-seq at the sequence level, together, these data suggest that m6A marks the majority of RNA:DNA hybrids in hPSCs. In addition, direct comparison of the efficiency of m6A and S9.6 antibodies revealed the superior sensitivity of m6A antibody in enrichment of m6A modified hybrids suggesting that the newly developed m6A DRIP offers a powerful indirect method for studying genome-wide distribution of RNA:DNA hybrids.

# Chapter 5

# Deposition of m6A on RNA:DNA hybrids correlate with transcription in a distinct chromatin environment

## 5. Deposition of m6A on RNA:DNA hybrids correlate with transcription

### 5.1. m6A and S9.6 peaks predominantly occur in transcribed genomic regions.

Consistent with co-transcriptional generation of R-loops (Huertas and Aguilera, 2003; Ginno et al., 2012; Stork et al., 2016), the majority of RNase H-sensitive m6A and S9.6 peaks were detected in transcribed regions of the genome (protein coding, anti-sense, and lncRNA coding regions, see Figure 4.6A). Therefore, the expression levels of the corresponding genes were assessed to understand if m6A and S9.6 peaks correlate with their transcriptional activity using total RNA-seq isolated from REBL-PAT hPSCs. In agreement with co-transcriptional generation of R-loops, transcribed genes were consistently over-represented (around 70%) in both m6A and S9.6 DRIP peaks datasets compared to genes that are transcriptionally silent (around 50%) in hPSCs (Figure 5.1A). Despite this, m6A peak-containing genomic sequences did not display significant enrichment for active, silent, bivalent or initiated promoters (Figure 5.1B). In addition, other than the fact m6A modified R-loops were mostly detected in transcribed genomic regions (Figure 4.4A), no significant correlation was found between m6A peak and the levels of transcriptional activity of the corresponding loci (Figure 5.1C-D). These results suggest that most R-loops are efficiently cleared from highly transcribed genomic regions.

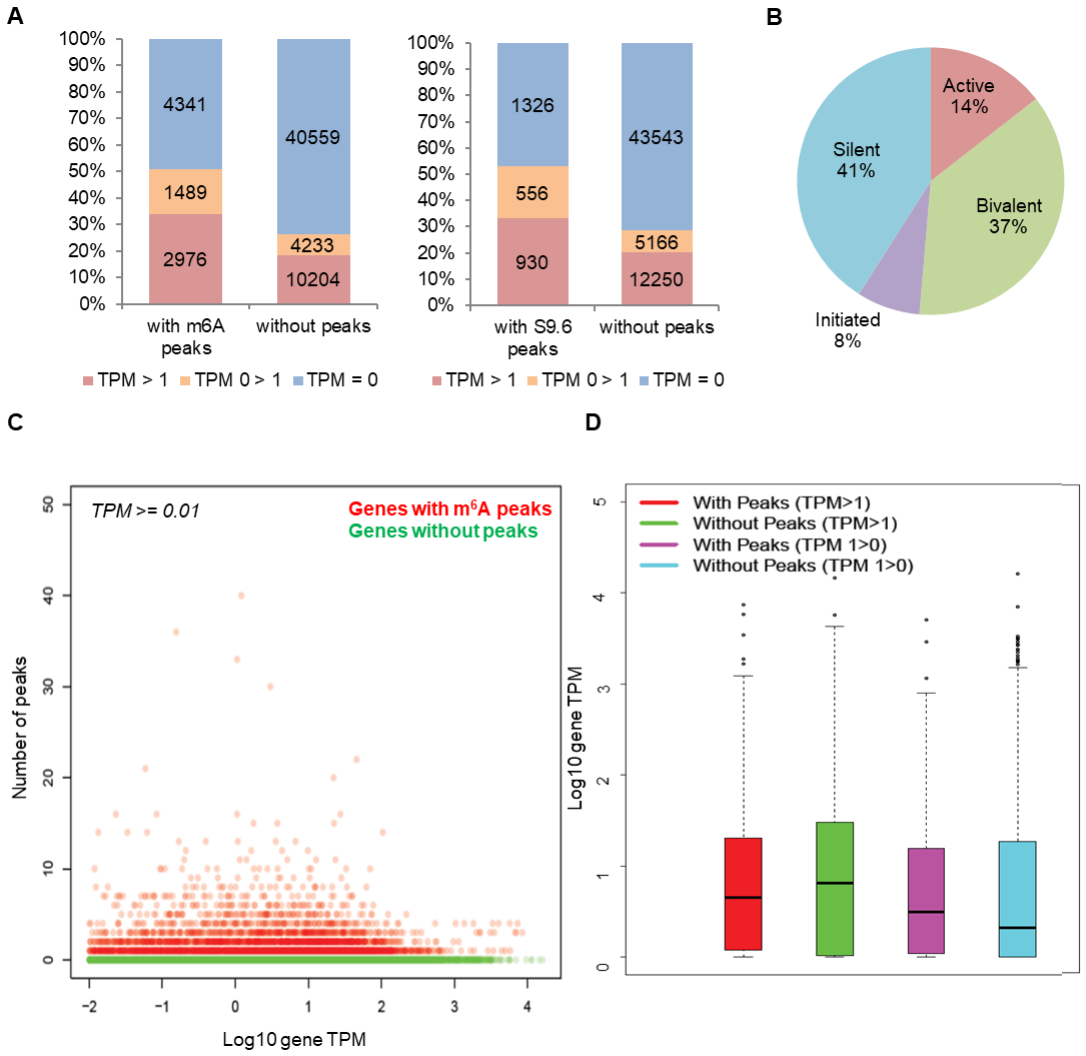



**Figure 5.1. The gene expression analysis of m6A and S9.6 peaks-containing genes.** (A) The ratios of genes with TPM (transcript per million reads) greater than 1 (TPM > 1), TPM more than 0 but less than 1 (0<TPM 1) and transcriptionally inactive genes (TPM = 0) amongst m6A or S9.6 narrow peaks-containing genes and genes without peaks in REBL-PAT hPSCs (B) Pie chart illustrating distribution of m6A DIP narrow peak-containing genes across their promoter categories (silent, active, bivalent and initiated) in hPSCs. (C) Distribution of Log10 gene TPM in relation to the number of m6A DIP narrow peaks in the gene for genes with TPM greater than 1 in REBL-PAT hiPSCs. (D) The box plots showing distribution of Log10 gene TPM for m6A narrow peaks-containing genes and genes without peaks with TPM greater than 1 (TPM > 1) and TPM more than 0 but less than 1 (0<TPM 1) The elements of the box plot are: center line, median; box limits, upper and lower quartiles; whiskers, 1.5× the interquartile range; n, sample size (genes); dots, outliers.

## 5.2. m6A and S9.6 DRIP-seq peaks strongly correlate with transcriptional activity of retroelements

Interestingly, the same assessment in retroelements revealed significant enrichment of m6A modified RNA:DNA hybrids in the open reading frame of evolutionary novel and transcriptionally active LINE1 (L1-Ta and L1-pre-Ta) retro-elements that comprise about one fifth of the human genome (Figure 5.2A-B) (Klawitter et al., 2016). Moreover, compared to m6A DRIP peak distribution in evolutionary younger L1PA2 peaks, evolutionary older and transcriptionally less active LINE1 retro-elements (e.g. L1PA5) showed significantly reduced enrichment of m6A peaks, suggesting m6A/S9.6 peaks correlate with transcriptional activity across LINE1 elements (Figure 5.2A). In agreement with this, Short INterspersed Elements (SINEs) from the Alu class that are transcribed by RNA polymerase III (pol-III) but rely on LINE1 elements for retro-transposition and comprise more than 11% of the genome (Batzer and Deininger, 2002) showed significant enrichment of m6A and S9.6 peaks (Figure 5.2A-B). Interestingly, unlike the bias of m6A and S9.6 peaks towards younger LINE1 elements, both evolutionary young Alu-Y and old Alu-S elements showed comparable enrichment of m6A peaks (Figure 2C). Thus, these data suggest that the presence of m6A on RNA:DNA hybrids correlates with actively transcribed regions of the human genome, in both RNA pol-II (LINE1s) and RNA pol-III (Alu) transcribed loci.

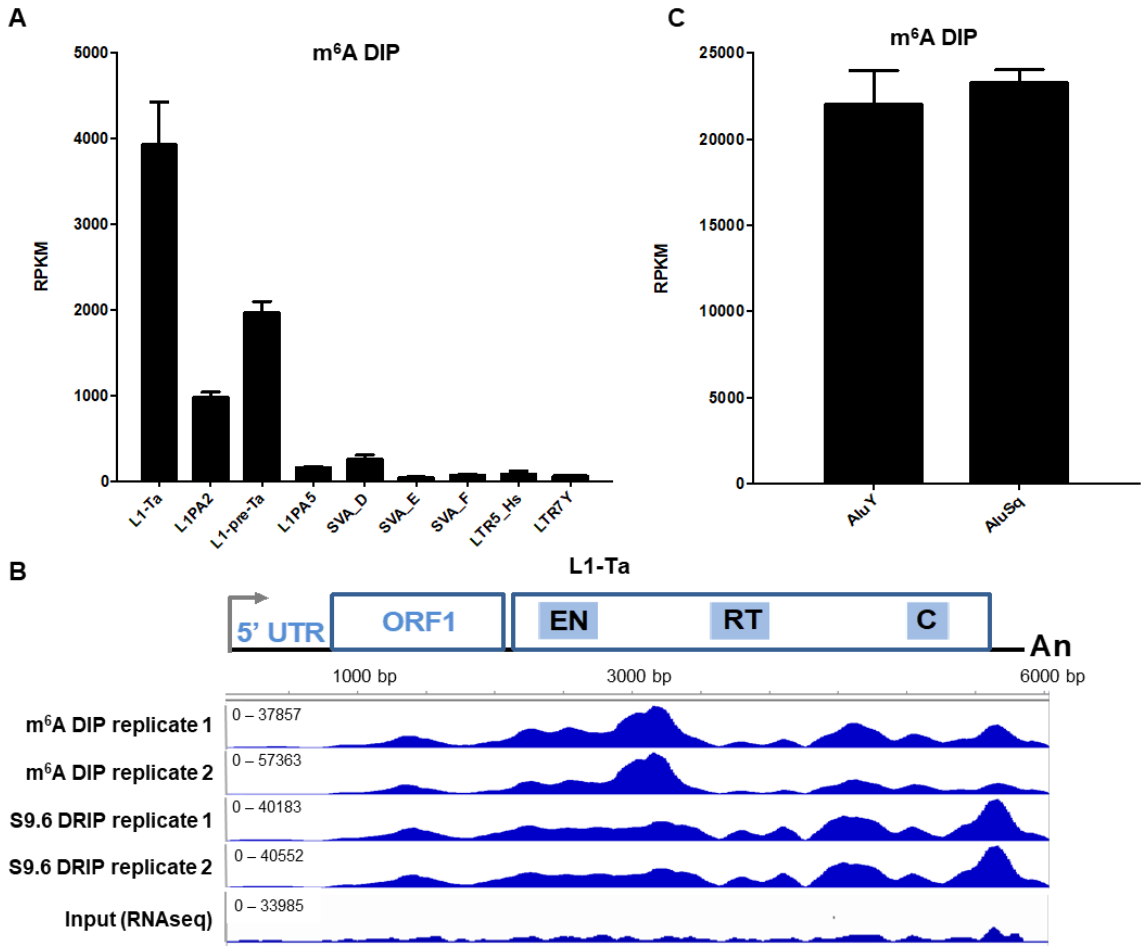



**Figure 5.2. The distribution of m6A marked RNA:DNA hybrids among different repetitive sequences**. (A) The distribution of normalized m6A DIP reads among different classes of LINE1, SVA and LTR repeats. (B) The distribution of normalized m6A DIP and S9.6 DRIP reads across L1-Ta transcript compared with the input (RNAseq results). (C) The distribution of normalized m6A DIP reads among the indicated classes of Alu repeats. Data are means ± SD, n=2 biologically independent samples (m6A DIP replicates) in (B, C).

### 5.3. m6A DIP and S9.6 DRIP peaks datasets do not overlap with majority of histone modifications in hPSCs.

Previously, co-transcriptional formation of R-loops has been suggested to contribute to chromatin patterning in regulatory regions (Ginno et al., 2012; Stork et al., 2016; García-Muse and Aguilera, 2019). To examine whether the presence of m6A/S9.6 peaks coincides with histone marks, Chromatin immunoprecipitation (ChIP)-derived datasets of histone modifications in hPSCs were aligned with m6A and S9.6 DRIP peaks. Unexpectedly, no significant correlation between any of the analysed histone modification and the distribution of

m6A and S9.6 peaks was identified (Figure 5.3). Despite this, both m6A and S9.6 peak datasets clustered together (Figure 5.3). Taken together, these results suggest the lack of overlap between RNA:DNA hybrids and major histone marks.

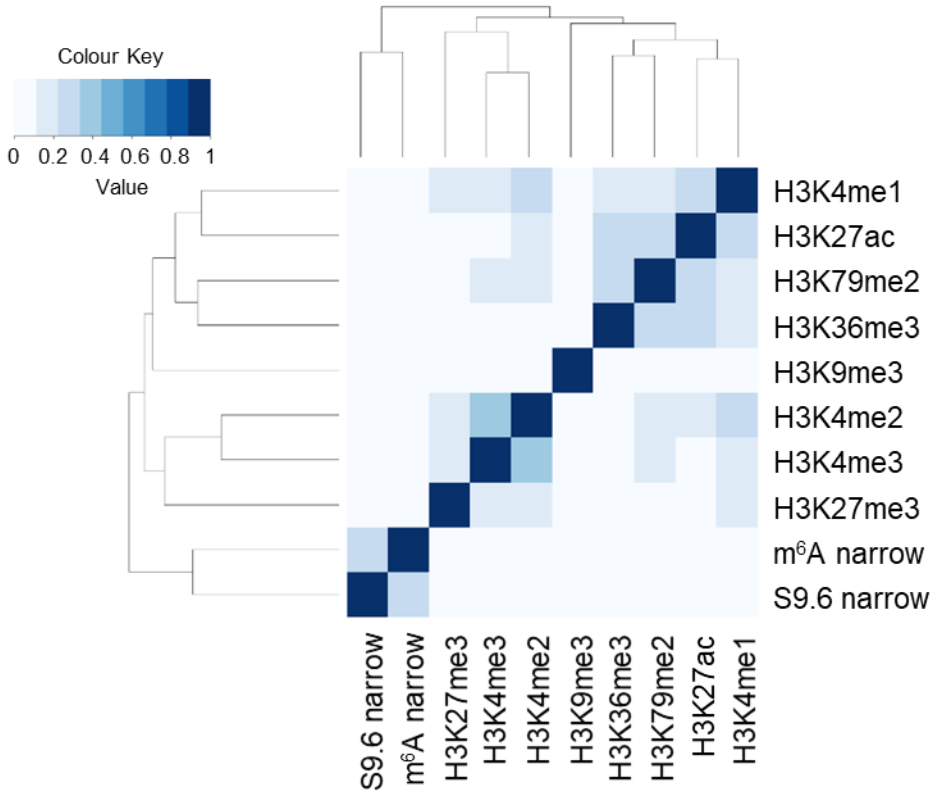



**Figure 5.3. Comparison of the m6A DIP and S9.6 DRIP peaks datasets with genomic distribution of histone modifications in hPSCs**. The results of clustering analysis of m6A DIP and S9.6 DRIP narrow peaks datasets and the indicated histone modifications' ChIP-seq peaks datasets (H3K4me1, GSM733782; H3K4me2, GSM733670; H3K4me3, GSM733657; H3K27me3, GSM733748; H3K27ac, GSM733718; H3K9me3, GSM1003585; H3K36me3, GSM733725; H3K79me2, GSM1003547).

### 5.4. Summary and conclusion

Overall, in line with co-transcriptional formation of RNA:DNA hybrids (Huertas and Aguilera, 2003; Stork et al., 2016), interspersed elements and individual loci showed m6A/S9.6 peaks are enriched in transcribed regions of the human genome. These results are in agreement with recent studies demonstrating co-transcriptional deposition of m6A on nascent RNAs by the

catalytic subunit of mammalian N6-adenosine-methyltransferase, METTL3 (He, 2010; Ping et al., 2014; Ke et al., 2017).

# Chapter 6

# m6A RNA:DNA hybrids exhibit cell cycle stage-specific dynamics in hPSCs

## 6. m6A RNA:DNA hybrids exhibit cell cycle stage-specific dynamics in hPSCs

Although studies implicating RNA:DNA hybrids in a myriad of cellular processes considerably expanded in recent years (Richard and Manley, 2017; Niehrs and Luke, 2020), little was known about the *dynamics* of RNA:DNA hybrids during the cell cycle. Nevertheless, mitotic chromatin of multiple human cell lines consistently showed RNase H sensitive high m6A immunostaining signal compared to interphase cells (Figures 3.4A), suggesting the rate of m6A deposition or degradation varied significantly depending on the cell-cycle stage.

### 6.1. Cell-cycle dependent m6A deposition on RNA:DNA hybrids in human cells.

To examine the cell cycle dynamics of m6A modified R-loops, parallel m6A and S9.6 DRIP were performed on FACS-sorted G0/G1, S and G2/M hPSCs populations (Figure 6.1.1A) and changes in R-loops levels corresponding to LINE1 elements (Figure 6.1.1B) and individual intronic regions (Figure 6.1.1C) were assessed for each cell-cycle phase using quantitative PCR (qPCR). In these experiments, both m6A and S9.6 immunoprecipitated RNA:DNA hybrids showed preferential accumulation in LINE1 regions during S phase, peak at G2/M and drastically decrease at G0/G1 phases of the cell cycle in hPSCs (Figure 6.1.2A-B). In contrast, α-satellite repeats showed no significant changes in both m6A and S9.6 retrieved RNA:DNA levels at any stage of hPSCs cell cycle. Since R-loops are predominantly co-transcriptionally formed (Huertas and Aguilera, 2003; Stork et al., 2016), these results implied that LINE1 repeats are also transcriptionally active during S and G2/M cell cycle phases in hPSCs. In agreement with this, a recent study reported that cells in S and mitotic phases of the cell cycle exhibit enrichment for the highly active LINE1 mRNAs (Mita et al., 2018). Similarly, m6A and S9.6 DRIP-qPCR show comparable accumulation of intronic R-loops during S and G2/M phases before significantly dropping in G0/G1 phase of the cell cycle in hPSCs (Figure 6.1.2C-D). Importantly, these cell cycle-specific changes in R-loop levels were essentially equivalent

in both m6A and S9.6 DRIP experiments, suggesting that m6A is present on RNA:DNA hybrids throughout all stages of the cell cycle (Figure 6.1.2A-D).

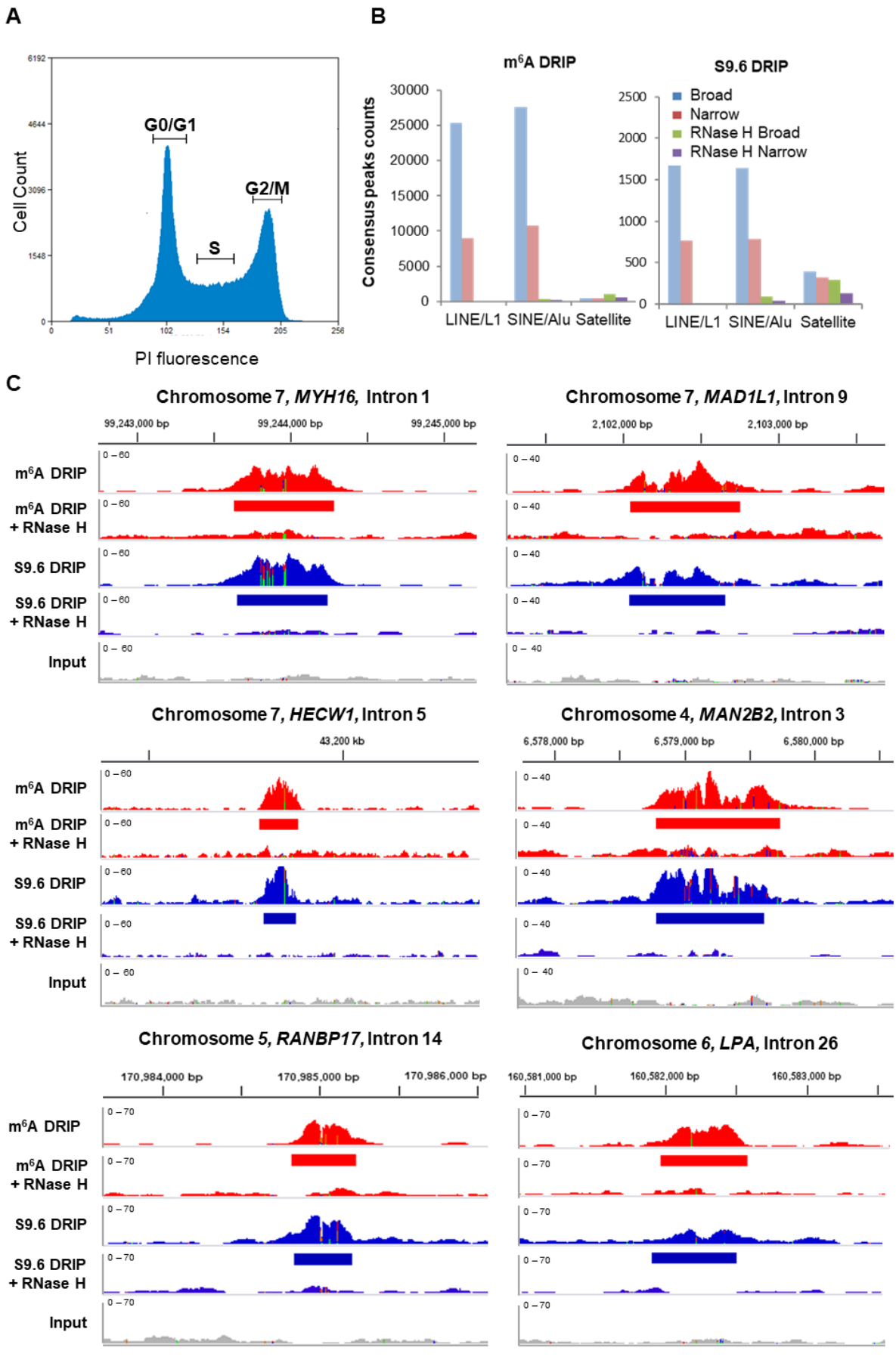

A
Cell Count
G0/G1
S
G2/M
PI fluorescence
B
m6A DRIP
S9.6 DRIP
Consensus peaks counts
Broad
Narrow
RNase H Broad
RNase H Narrow
LINE/L1
SINE/Alu
Satellite
C
Chromosome 7, MYH16, Intron 1
Chromosome 7, MAD1L1, Intron 9
Chromosome 7, HECW1, Intron 5
Chromosome 4, MAN2B2, Intron 3
Chromosome 5, RANBP17, Intron 14
Chromosome 6, LPA, Intron 26
m6A DRIP
m6A DRIP + RNase H
S9.6 DRIP
S9.6 DRIP + RNase H
Input

**Figure 6.1.1. FACS sorting of hPSCs and selection of RNA:DNA hybrid containing loci for analysis of their cycle-cycle**. (A) The diagram illustrating gating of single hiPSCs using PI-Area and PI-Width signals (left panel) and DNA content frequency histogram (right panel) of a representative hiPSC population used for cell cycle analysis. hPSCs at G0/G1, S and G2/M phases are marked. (B) m6A DIP and S9.6 DRIP consensus narrow peak counts of the indicated repetitive elements in hPSCs (C) The coverage plots of m6A DIP and S9.6 DRIP densities (CPK) in the intronic regions of the indicated genes. m6A and S9.6 peaks are marked with red and blue rectangles.

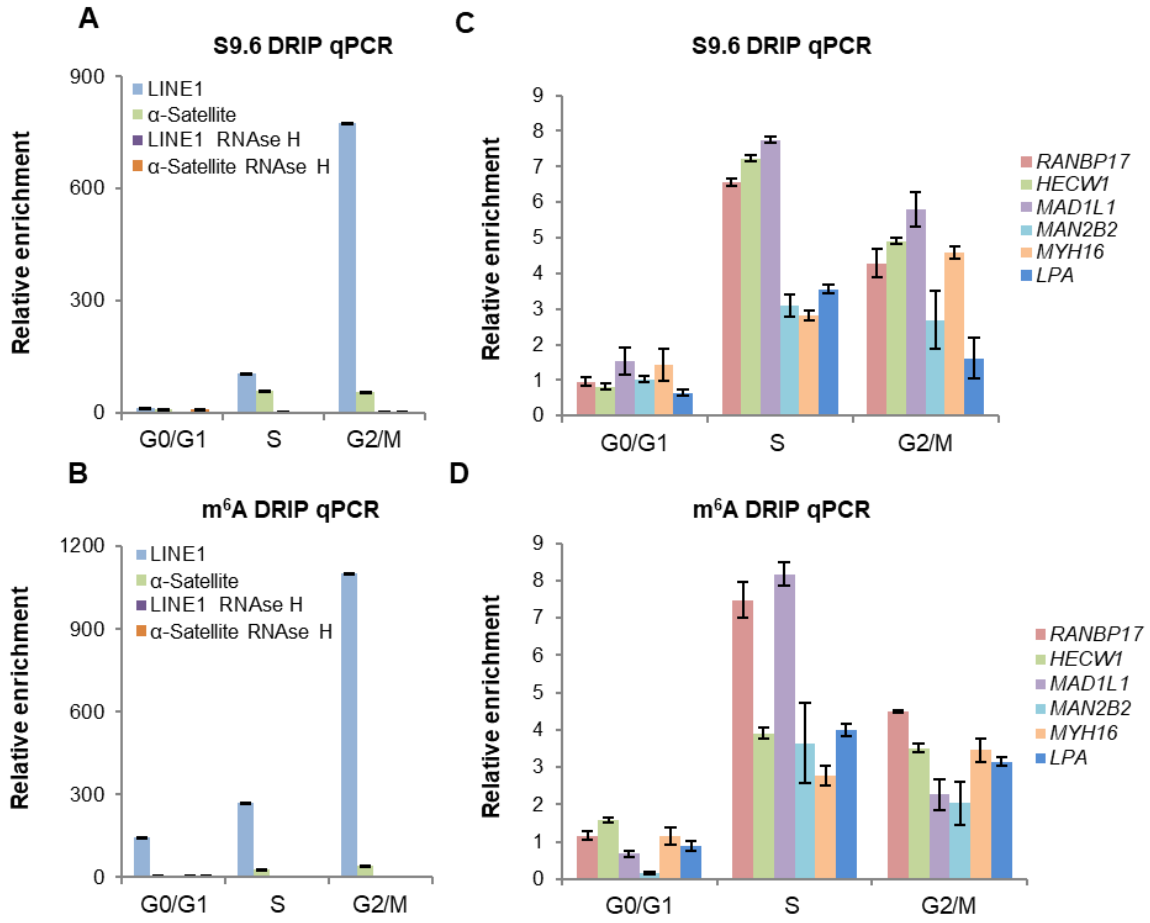



**Figure 6.1.2. Cell-cycle dependent dynamics of RNA:DNA hybrids**
(A-B) Results of m6A DIP and S9.6 DRIP–qPCR of the indicated repeats performed on FACS sorted hPSCs at different cell cycle phases in presence and absence of RNase H. Generic primers amplifying evolutionarily young L1Hs were used. Results of S9.6 DRIP (C) and m6A DIP–qPCR (D) of RNA:DNA peaks localized in the introns of the indicated genes (see figure 6.1B) performed on FACS sorted hPSCs. A-D, Data are means ± s.d., n= 3 independent experiments.

## 6.2. Depletion of nuclear RNase H1 results in accumulation of m6A marked RNA:DNA hybrids in hPSCs

Sensitivity to RNase H is one of the defining features of R-loop structures (Cerritelli and Crouch, 2009). Consistent with this, small interfering RNA (siRNA)-mediated 60-70% depletion of nuclear RNase H1 transcripts in hPSCs (Figure 6.2A) resulted in accumulation of m6A- and S9.6-immunoprecipitated RNA:DNA hybrids on repetitive and intronic loci (Figure 6.2B) without significant changes in LINE1 transcription (Figure 6.2C). Confirming the specificity of RNase H1 depletion to m6A and S9.6 DRIP peak containing loci, no enrichment for m6A- or S9.6 DRIP-qPCR was detected for genomic regions downstream of m6A and S9.6 peak sites (Figure 6.2B,D-E). Taken together, these results revealed that the formation of RNase H sensitive m6A-marked RNA:DNA hybrids occurred *in vivo* which is consistent with the co-transcriptional deposition of m6A (Ke et al., 2017), as well as the generation of RNA:DNA hybrids (Huertas and Aguilera, 2003; Stork et al., 2016).

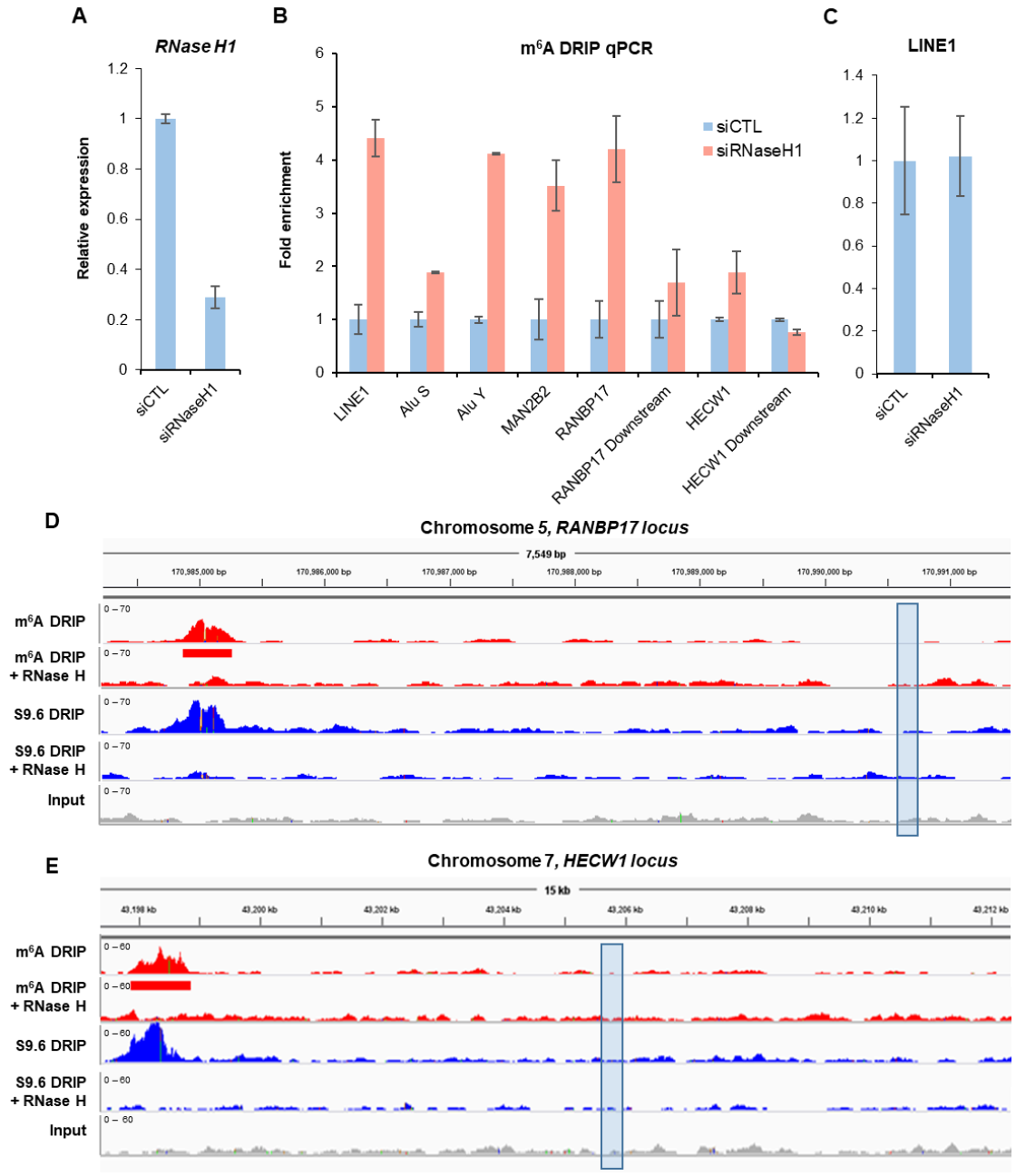



**Figure 6.2. m6A DIP signal is increased upon RNase H1 knockdown in hPSCs**.
(A) Relative expression of RNase H1 RNA in hPSCs transfected with control non-targeting (siCTL) and RNase H1 (siRNaseH1) siRNAs. (B) The results of m6A DIP qPCR of the indicated repeats and intronic sequences performed on siCTL and siRNaseH1 hiPSCs. The regions without peaks (RANBP17 Downstream and HECW1 Downstream) were used as controls. Generic primers amplifying Alu elements from the indicated families and evolutionarily young L1Hs were used for DRIP qPCRs and qPCR. Data are means ± SD, n = 3 independent experiments. (C) Relative expression of LINE1 RNA in hPSCs transfected with control non-targeting (siCTL) and RNase H1 (siRNaseH1) siRNAs. (D) The coverage plots of m6A and S9.6 DRIP/DIP densities (CPK) in the regions located downstream of the m6A and S9.6 peaks (marked with red rectangles). The location of regions that were used as controls in (B) is designated by blue rectangles.

### 6.3. Sequential S9.6 and m6A DRIP confirm enrichment of R-loops across multiple loci.

To further confirm m6A marks presence on S9.6 immunoprecipitated fragments, two sequential IP reactions were performed, S9.6 DRIP followed by m6A DIP-qPCR (Figure 6.3A). Consistent with previous results, individual intronic (Figure 6.3B) and repetitive sequences (Figure 6.3C) showed enrichment in the two-round IP procedure, confirming the presence of m6A on S9.6 retrieved R-loops.

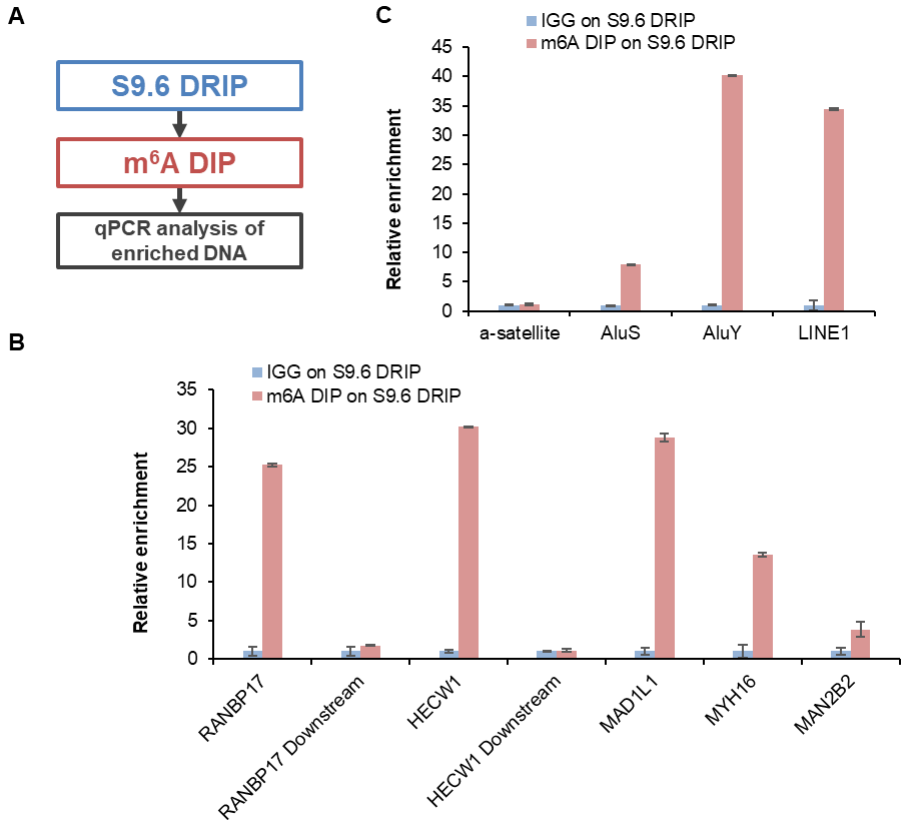



**Figure 6.3. m6A is detectable on the RNA components of R-loops.** (A) Schematic representation of the two rounds IP procedure, S9.6 DRIP followed by m6A DIP procedure. (B, C) The results of the two round (S9.6 DRIP followed by m6A DIP) DRIP/ DIP performed on individual intronic (b) and repetitive (c) m6A/S9.6 peak containing sequences. Representative results of analysis of one of 3 independent biological samples are presented. Data are means ± SD, n = 3 technical repeats.

### 6.4. Summary and conclusion

Taken together, these results confirm the presence of m6A on conventional S9.6 retrieved RNA:DNA hybrids and point to a model wherein the cell cycle dictates the rates of R-loop deposition and clearance.

# Chapter 7

# m6A regulates stability of RNA:DNA hybrids in human cells

## 7. m6A regulates the stability of RNA:DNA hybrids in human cells

Shortly after the discovery of m6A (Desrosiers et al., 1974; Perry and Kelley, 1975), seminal studies of m6A function demonstrated that its addition to nascent pre-mRNAs results in instability of target mRNAs (Sommer et al., 1978; Ke et al., 2017). This was shortly followed up by the cloning of m6A writer METTL3 (Bokar et al., 1997) whose depletion was shown to increase the average half-lives of target mRNAs (Sommer et al., 1978; Ke et al., 2017). The demonstration that m6A marks the majority of RNA:DNA hybrids (Figure 4.3.2A), raised the interesting hypothesis of whether m6A, similar to the consequence of its presence on mRNAs (Wang et al., 2014), can also regulate the stability of RNA:DNA hybrids.

### 7.1. METTL3 depleted hPSCs exhibit increased accumulation of R-loops.

To investigate whether m6A has a destabilising effect on R-loops, the m6A writer enzyme METTL3 (Bokar et al., 1997; Bokar, 2005) was depleted in hPSCs using siRNA treatment (Figure 7.1A). Although this did not significantly change the expression level of R-loop-containing genomic loci (Figure 6.4A-B), compared to untreated control samples, METTL3 depleted hPSCs showed increased cell-cycle dependent accumulation of R-loops in intronic sequences (Figure 6.4C), LINE1 (Figure 6.4D), and Alu sequences (Figure 6.4E). These results strongly suggest the involvement of METTL3 deposited m6A in regulating the stability of RNA:DNA hybrids.

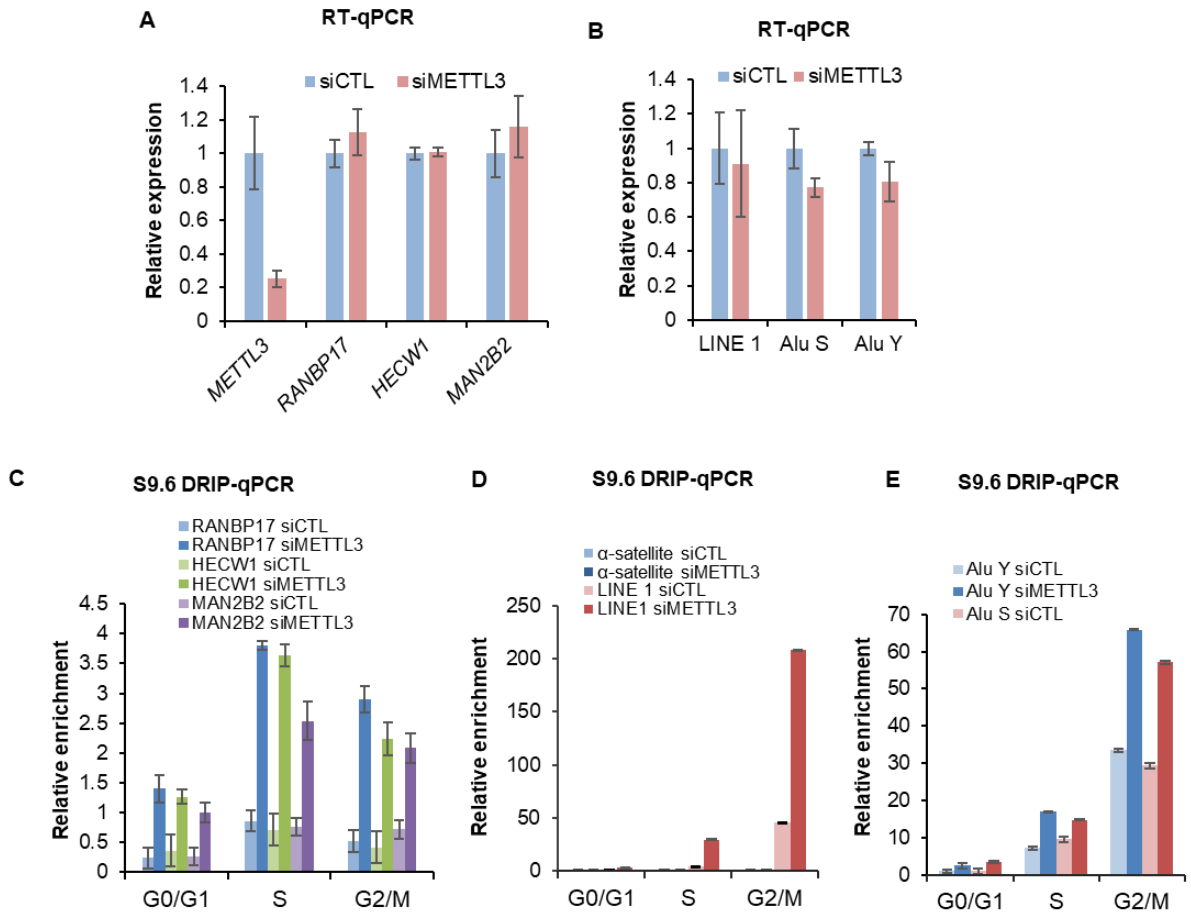



**Figure 7.1. METTL3 depletion leads to accumulation of RNA:DNA hybrids in hPSCs.** METTL3 depletion leads to accumulation of RNA:DNA hybrids in hPSCs. (a) The results of S9.6 DRIP qPCR of the indicated repeats and intronic sequences performed on siCTL and siMETTL3 hiPSCs sorted at different cell cycle phases. (b) Relative expression of the indicated transcripts in siCTL and siMETTL3 hiPSCs. The difference between the values of Y-axes in (a) and in Fig. 3b is due to incorporation of RNase H control samples in the analysis shown in Fig. 3b. Data are means ± SD, n = 3 independent experiments.

## 7.2. METTL3 depleted hPSCs display genome-wide increase in m6A marked R-loop levels.

Highlighting the importance of m6A in regulating the stability of R-loops, as little as 29% reduction in m6A brought about by METTL3 depletion in hPSCs transcriptome (Figure 7.2A) led to 35% increase in m6A retrieved from RNase H genomic DNA isolated from METTL3 depleted hPSCs (Figure 7.2B). These results provide additional evidence that modulation of m6A levels directly influences the stability of R-loops. Correspondingly, compared to untreated cells, hPSCs treated with siMETTL3 showed 3.8-fold increase in the

numbers of consensus m6A DRIP peaks (Figure 7.2C – D). Importantly, the presence of siMETTL3-specific m6A peaks did not correlate with changes in transcription of the corresponding genes (Figure 7.2D-E).

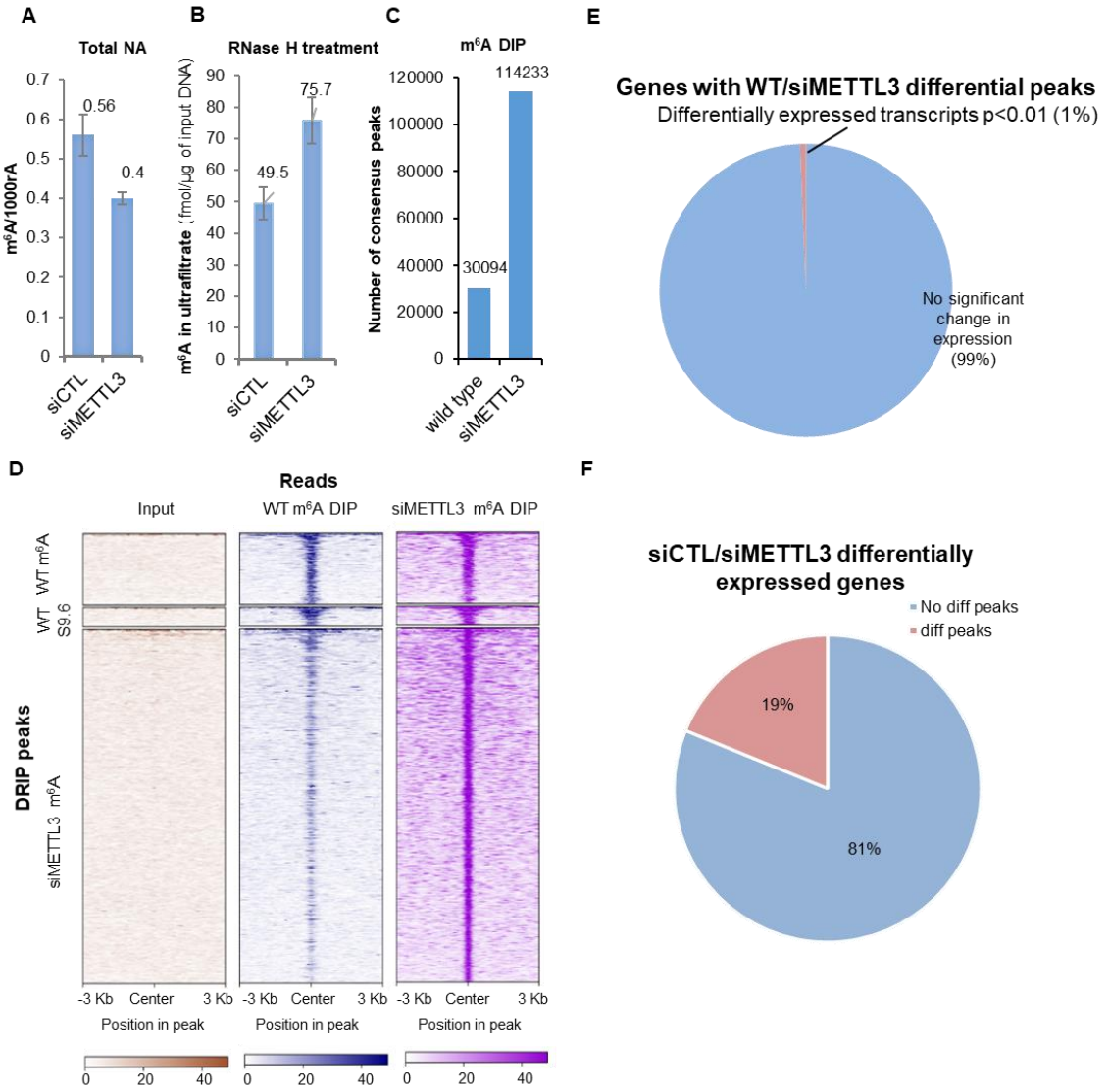



**Figure 7.2. METTL3 depletion leads to an increase in the numbers of m$^6$A DIP peaks in hPSCs. (**A) SID-UPLC-MS/MS quantification of m$^6$A in the total nucleic acids isolated from siCTL and siMETTL3 hPSCs. (B) SID-UPLC-MS/MS quantification of m$^6$A retrieved from genomic DNA in siCTL and siMETTL3 hPSCs following RNase H treatment. The data are means of 2 experiments ± SD. (C) The total numbers of consensus peaks of m$^6$A DIP performed on wild type and siMETTL3 hiPSCs. (D) Heatmaps showing the distribution of density of indicated reads across genomic regions containing peaks (3 kb around peak center) obtained in three experiments: S9.6 DRIP performed on wild type (WT) hPSCs, m$^6$A DIP carried out on WT cells and m$^6$A DIP performed on siMETTL3 hPSCs. The colour of each line represents the density of reads for a given peak. The width of the heatmaps is normalized by peak length. (E) Pie chart showing the percentages of transcripts differentially expressed between siCTL and siMETTL3 hPSCs (p<0.01) and transcripts without significant changes

in expression amongst genes containing $m^6A$ DIP peaks specific only for siMETTL3 hPSCs. (F) Pie chart demonstrating the percentages of genes containing $m^6A$ DIP peaks specific only for siMETTL3 hPSCs (diff peaks) and genes without such peaks (No diff peaks) corresponding to the transcripts differentially expressed between siCTL and siMETTL3 hPSCs.

### 7.3. m6A writer and readers are amongst proteins physically interacting with R-loops

The identification of m6A binding proteins implicated m6A in almost all steps of RNA metabolism (Meyer and Jaffrey, 2014; Hsu et al., 2017). Broadly, m6A reader proteins can be classified as YTH-domain containing proteins (YTHDF1, YTHDF2, YTHDF3, YTHDC1 and YTHDC2) (Lasman et al., 2020) and non-YTH domain m6A readers with the nuclear m6A reader HNRNPA2B1 provides an example of one of these (Alarcón et al., 2015). Whereas YTHDF1 and YTHDF3 was shown to regulate the translation efficiency of m6A containing mRNAs (Wang et al., 2015; Shi et al., 2018), YTHDF2 has been shown to promote degradation of m6A containing mRNAs (Wang et al., 2014). These findings raised the intriguing hypothesis that these reader proteins may interact with R-loops via m6A. In support of this hypothesis, the m6A writer, METTL3 along with YTHDF1, YTHDF2 and HNRNPA2B1 were detected by slot-blot amongst proteins bound to S9.6 retrieved R-loops (Figure 7.3). These results suggest the involvement of m6A reader proteins in mediating m6A function in R-loops.

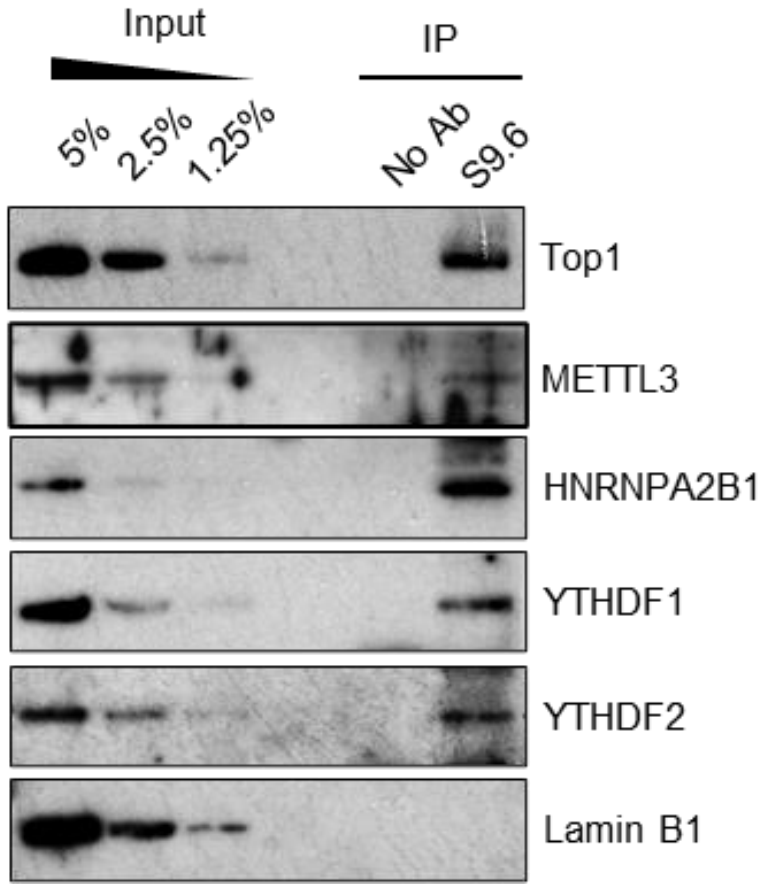



**Figure 7.3. m6A writer and reader proteins interact with RNA:DNA hybrids**.
Immunoblot of RNA:DNA hybrid protein co-immunoprecipitation (IP)-probed with indicated antibodies. Top1 and Lamin B1 serve as positive and negative controls, respectively, for R-loop immunoprecipitation. The experiments were repeated independently two times for METTL3 and three times for other proteins, with similar results.

### 7.4. Summary and conclusions

Taken together, these findings suggest that METTL3-depletion-mediated, reduced m6A deposition led in turn to reduced m6A turn-over rates, resulting ultimately in R-loop accumulation. Moreover, the discovery of physical interaction of m6A reader proteins with native R-loops could provide a physiological rationale for m6A deposition on R-loops in the first place.

# Chapter 8
# Dissecting the nature of key m6A reader protein, YTHDF2, interaction with m6A modified RNA:DNA hybrids

## 8. Dissecting the nature of YTHDF2 interaction with m6A modified RNA:DNA hybrids.

### 8.1. Cytoplasmic YTHDF2 localizes to chromatin of mitotic hPSCs

Though the detection of m6A reader proteins on S9.6 immuno-precipitated R-loops (see section 7.3) suggested the potential involvement of these reader proteins in R-loop biology (Figure 7.3), it was not clear whether any of these interactions is functional. Moreover, since the interactions were examined in R-loops retrieved from asynchronous cells, it was unclear whether any of these reader proteins interact with mitotic chromatin, enriched in m6A modified R-loops. To examine the localisation of m6A binding proteins, first, m6A and HNRNPA2B1 co-immunostaining was performed. While m6A staining exhibited the characteristic high signal in mitotic chromatin, HNRNPA2B1 staining was specifically excluded from the chromatin in mitotic cells (Figure 8.1A-B). Next, the YTH-domain containing proteins were examined for potential interaction with the R-loops enriched mitotic chromatin. YTHDF1 co-immunostaining with the mitotic cell marker, Phospho-H3 showed that YTHDF1 exhibits predominantly cytoplasmic localization in both interphase and mitotic hPSCs (Figure C-D). Although YTHDF2 staining signal was predominantly cytoplasmic in interphase cells, curiously, YTHDF2 exhibited strong co-localisation with Phospho-H3 in mitotic chromatin (Figure 8.1E-G). These results suggest a potential interaction between m6A enriched R-loops in mitotic cells and YTHDF2.

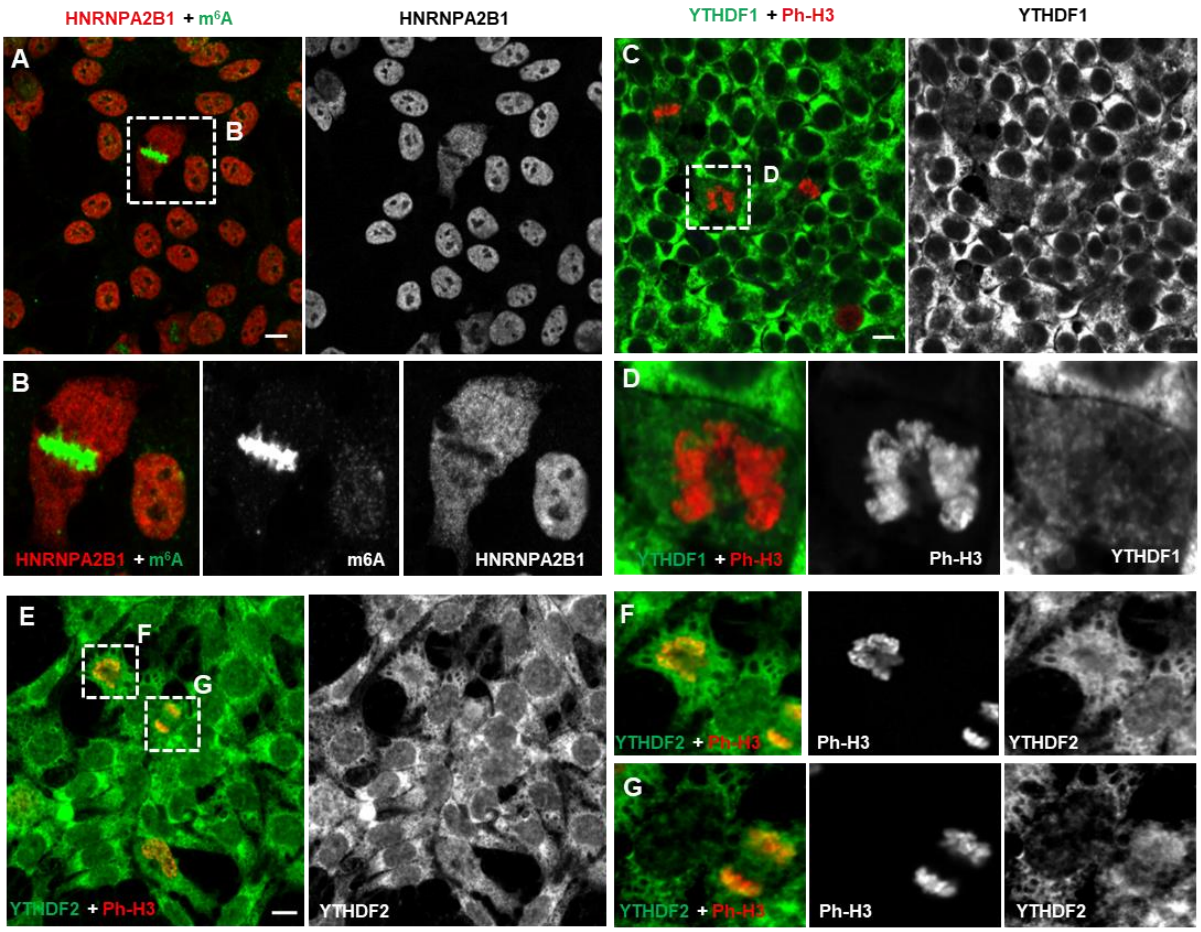



**Figure 8.1. Immunostaining of hPSCs for m6A reader proteins**. (A, B) Immunostaining of hiPSCs for HNRNPA2B1 and m6 A (C,D) Immunostaining of hiPSCs by using anti-YTHDF1 and anti-pH-H3 antibodies imaged at two different magnifications. imaged at two different magnifications. (E,F,G) Immunostaining of hiPSCs for YTHDF2 and pH-H3 imaged at two different magnifications. Merged views and YTHDF1/HNRNPA2B1/YTHDF2 channels (A-G). Locations of the views shown in B,D,F,G are marked with dashed rectangles in A,C,E. All scale bars, 10 μm. Experiments shown in A-G were repeated independently four times with similar results.

### 8.2. YTHDF2 co-localises with R-loops in *in vivo*

Since levels of m6A modified R-loops peaked in mitotic cells, the localisation of YTHDF2 suggested the localisation of YTHDF2 to R-loops. To examine this, YTHDF2 co-staining with the S9.6 R-loops antibody was performed in fixed hPSCs. Interestingly, YTHDF2 co-staining with S9.6 (Figure 8.2A) revealed a high degree of YTHDF2 staining signal co-localization with R-loops (Figure 8.2B). Moreover, YTHDF2 and S9.6 signals showed more than 90% overlap across hPSCs nuclei. Although these results revealed the co-localisation of YTHDF2 and RNA:DNA hybrids, it was not clear whether this interaction is direct or m6A dependent.

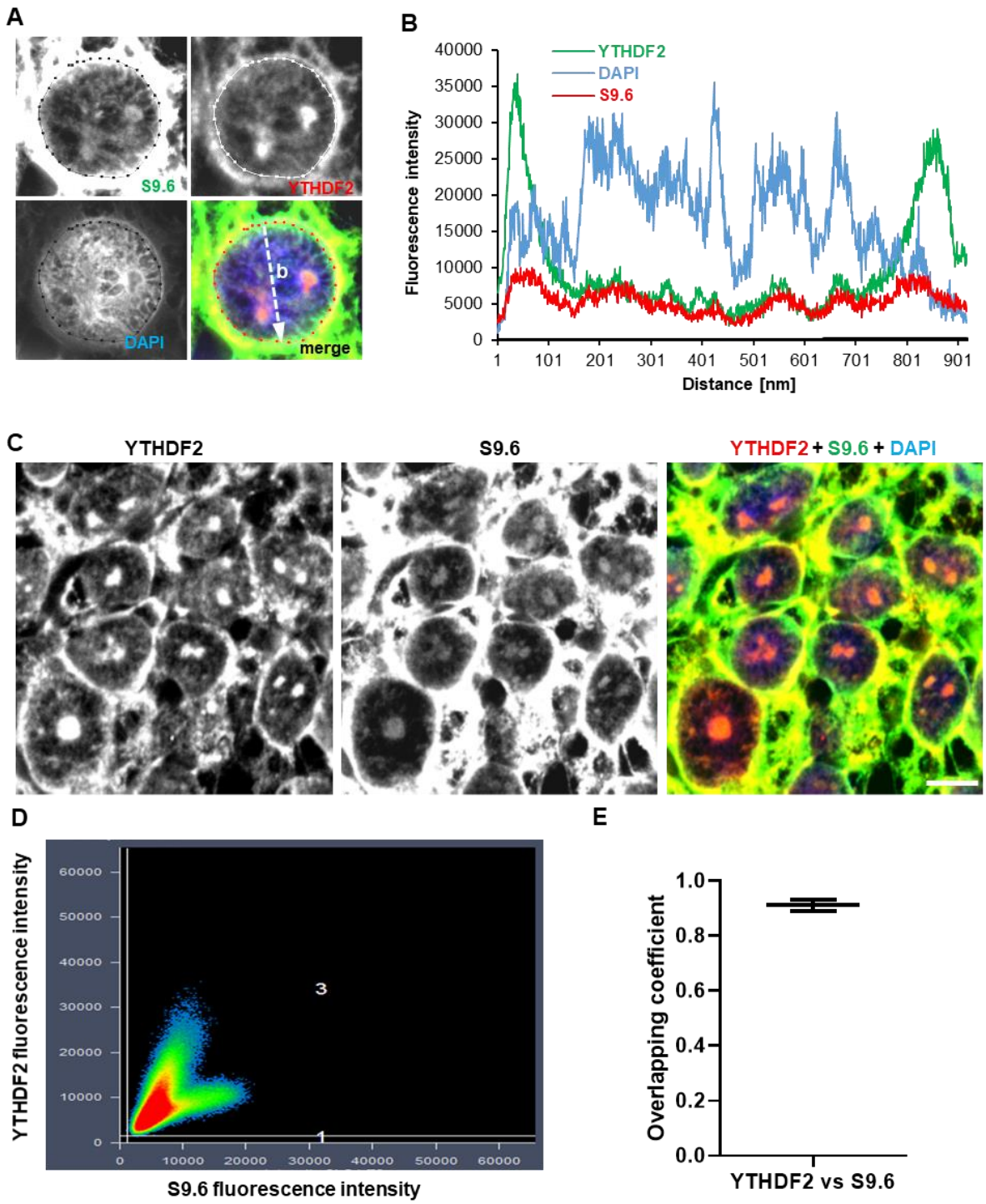



**Figure 8.2. YTHDF2 co-localizes with R-loops *in vivo***. YTHDF2 co-localizes with R-loops in vivo and interacts with synthetic m6A-marked RNA:DNA hybrids in vitro. (A) Co-immunostaining of YTHDF2 with R-loops (S9.6) in a representative hPSCs interphase nucleus. Merged view and individual channels are shown. The area used for S9.6/YTHDF2 signals quantification is indicated. The arrow designates the region used for generation of signal intensity profile shown in (B). (b) The profile showing intensities of YTHDF2, S9.6 and 4,6-diamidino-2-phenylindole (DAPI) signals across the nuclear region marked with an arrow in (A). (C) Representative image of hPSCs immunostained for YTHDF2 and R-loops (S9.6). Merged view and YTHDF2/S9.6 channels are shown. Scale bar = 20 μm The

experiments shown in (A, C) were repeated independently 3 times with similar results. (D) Scatter diagram for YTHDF2 and S9.6 fluorescence intensities in an individual hPSCs nucleus. (E) The value of overlapping coefficient of YTHDF2 vs S9.6 intensities quantified for 20 cells immunostained for YTHDF2 and S9.6. The centre value is median, error bars show minimal and maximal values.

### 8.3. YTHDF2 interaction with synthetic RNA:DNA hybrids is m6A dependent

In order to examine whether the interaction of YTHDF2 with R-loops is mediated by the presence of m6A on the RNA, the binding affinity of YTHDF2 to γ32P-ATP labelled 42-mer synthetic hybrids and single stranded RNAs modified with and without m6A in the consensus GGAC motif was examined using in electrophoretic mobility-shift assays (EMSA). Consistent with previous observations (Wang et al., 2014) human recombinant YTHDF2 exhibited higher binding affinity towards m6A containing single stranded RNA compared to oligoes with no m6A modification. Similarly, YTHDF2 exhibited higher binding affinity towards m6A-containing RNA hybridized to DNA compared to RNA:DNA hybrids without m6A, both in low and high salt conditions (Figure 8.3A-B). These results confirm the interaction of YTHDF2 with RNA:DNA hybrids and reveal essential role for m6A in mediating this interaction.

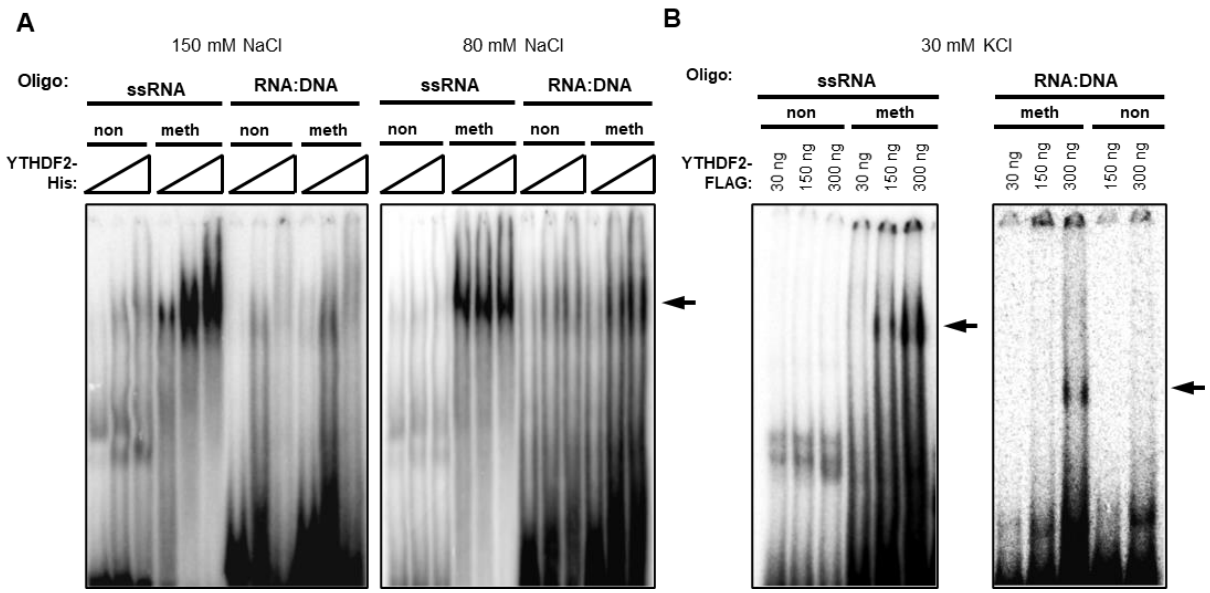



**Figure 8.3. YTHDF2 interacts with synthetic m6A-marked RNA:DNA hybrids *in vitro*.** (A) The results of EMSA using recombinant YTHDF2-His fusion (B) or YTHDF2-FLAG (B) and 0.15 pmol of unmodified (non) or m6A-containing (meth) RNA oligonucleotides (ssRNA) or corresponding synthetic RNA:DNA hybrids (RNA:DNA). Triangles indicate increasing concentrations of the protein (10, 100, 300 ng). Concentrations of the recombinant protein (B) and NaCl/KCl in the binding buffer are indicated. The RNA and RNA:DNA-YTHDF2 protein complexes are arrowed.

### 8.4. YTHDF2 shows higher binding affinity towards m6A containing RNA:DNA hybrids compared to unmodified RNA:DNA hybrids

To further clarify the preferential substrate specificity of YTHDF2 towards m6A modified RNA and RNA:DNA hybrids, its binding affinity towards methylated and non-methylated synthetic RNA and RNA:DNA hybrids was examined in microscale thermophoresis (MST) analysis. These data showed that YTHDF2 exhibited similar binding affinity towards m6A on single stranded RNA (Kd:15.5 +/-4.8 μM) or RNA:DNA hybrids (Kd:16.2 +/-5.0 μM). Confirming YTHDF2 interaction with RNA and RNA:DNA hybrids dependency on m6A, YTHDF2 exhibited significantly lower binding affinity towards single stranded RNA and RNA:DNA hybrids that did not have m6A (Figure 8.4A-B). These results further confirm that binding of YTHDF2 to RNA:DNA hybrids is m6A dependent.

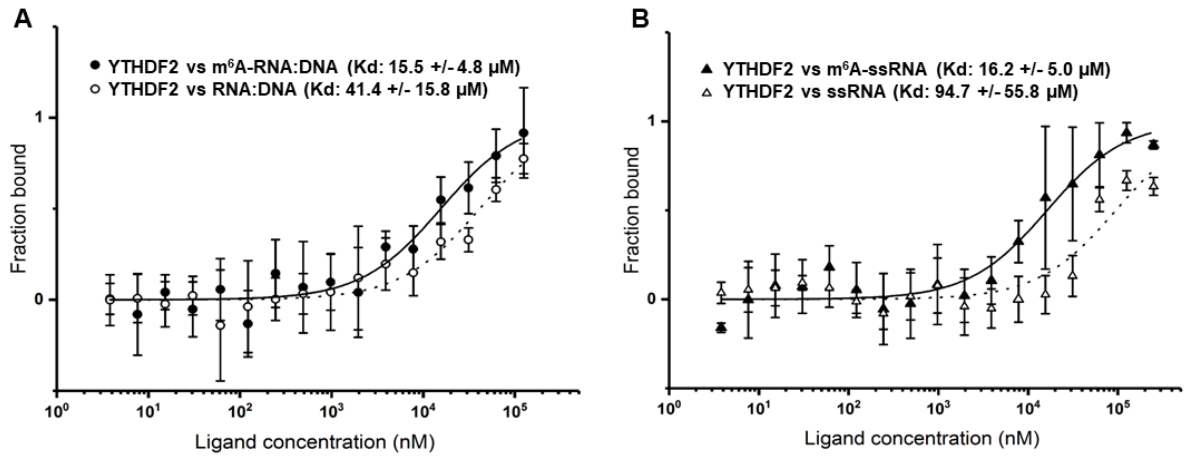



**Figure 8.4. YTHDF2 exhibit higher binding affinity towards m6A containing RNA:DNA hybrids compared with non-modified RNA:DNA duplexes**. Microscale thermophoresis binding curves for YTHDF2 interaction with m6 A-containing/non-modified RNA:DNA hybrid and m6A-marked/non-modified ssRNA synthetic substrates. Binding is shown as fraction of protein bound as a function of substrate concentration. Binding curves are fitted to the data points from experiments for m6 A-containing (filled circles/triangles) and unmodified (open circles/triangles) substrates. Dissociation constant (Kd) values are shown for each of the interactions. Error bars show s.d., the center values are means, n= 6 independent series of experiments.

### 8.5. Chromatin immunoprecipitation (ChIP) reveal that YTHDF2 intercats with R-loops containing genomic-loci *in vivo*.

To interrogate whether the interaction of YTHDF2 to m6A modified synthetic hybrids occurs *in vivo*, YTHDF2-bound chromatin was examined using ChIP-qPCR. YTHDF2 ChIP-qPCR results revealed recruitment of YTHDF2 to LINE1 elements and intronic sequences enriched in R-loops (Figure 8.5A). In contrast, ChIP-qPCR analysis of another m6A reader, HNRNPA2B1 revealed no binding of HNRNPA2B1 to LINE1 repeats even though the presence of HNRNPA2B1 at intronic sequences enriched in R-loops was detected in these experiments (Figure 8.5B). The binding of YTHDF2 to LINE1 elements and intronic

sequences enriched in R-loops suggests that YTHDF2 and R-loops interact *in vivo*.

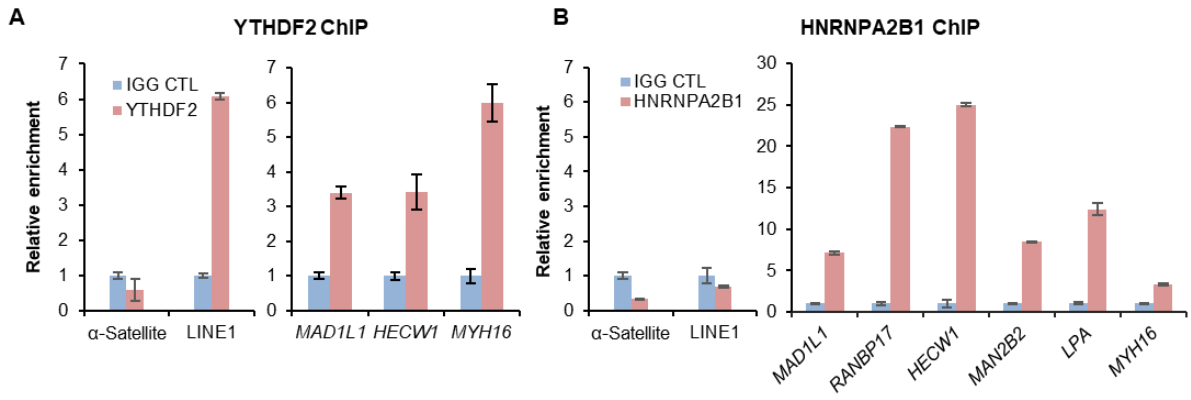



**Figure 8.5. ChIP-qPCR reveals YTHDF2 interacts with R-loop-containing loci *in vivo*.** (A-B) The results of YTHDF2 (A) and HNRNPA2B1 (B) ChIP-qPCR of the indicated repeats and intronic sequences performed on REBL-PAT hPSCs (A-B. Representative results of one of 3 independent experiments ± SEM are presented.

### 8.6. Enrichment of YTHDF2 at R-loop-containing genomic-loci in METTL3 and RNase H1 depleted hPSCs.

Since the ribonuclease RNase H1 mediate removal of cellular R-loops (Cerritelli and Crouch, 2009), numerous studies used RNase H1 depletion to study consequences of R-loop formation. The binding of YTHDF2 to R-loop enriched genomic loci raised the hypothesis that YTHDF2 is recruited as a response to R-loop formation. To test this and further verify YTHDF2 interaction with R-loops *in vivo*, RNase H1 depleted and control hPSCs were examined for YTHDF2 recruitment to chromatin. Analysis of YTHDF-2 ChIP-qPCR revealed that, unlike in untreated control cells, siRNase H1 depleted hPSCs show increased recruitment of YTHDF2 to both LINE1 repeats and intronic sequences following transient RNase H1 depletion (Figure 8.6A). Likewise, consistent with recruitment of HNRNPA2B1 to intronic sequences enriched in R-loops (Figure 8.5B), compared to untreated (control) cells, HNRNPA2B1 ChIP-qPCR analysis of RNase H1 depleted hPSCs revealed increased binding of HNRNPA2B1 to intronic sequences (Figure 8.6B). Next, YTHDF2 and HNRNPA2B1 ChIP-qPCR experiments in siMETTL3 depleted and untreated control hPSCs were performed to verify whether the *in vivo* interaction of YTHDF2 and HNRNPA2B1 with R-loops is m6A dependent. Consistent with

m6A dependent recruitment of both YTHDF2 and HNRNP2B1 to R-loops, depletion of METTL3 led to decreased binding of both m6A readers to R-loops containing genomic loci (Figure 8.6C-D). In sum, these results indicated the m6A depended binding of YTHDF2 and HNRNPA2B1 to R-loops and suggested functional relevance of these interactions.

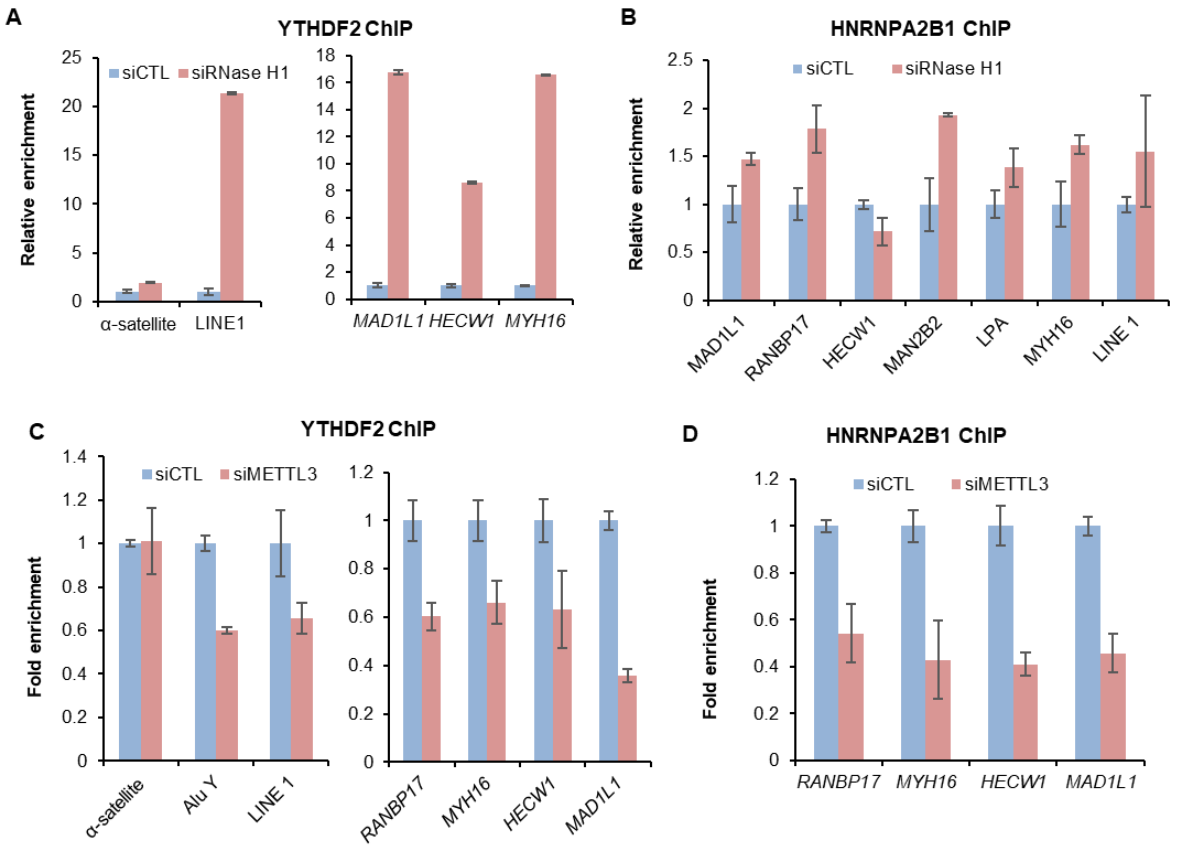



**Figure 8.6. Increased YTHDF2 recruitment to R-loop-containing genomic loci in RNase H1 or METTL3 depleted hPSCs**. The results of YTHDF2 (A,C) and HNRNPA2B1 (B,D) ChIP-qPCR of the indicated repeats and intronic sequences performed on siCNTL and siRNase H1 hPSCs (A,B) or on siCTL and siMETTL3 hPSCs (C,D). Representative results of one of 3 independent experiments ± SEM are presented.

### 8.7. YTHDF2 depletion leads to accumulation of RNA:DNA hybrids in human cells

To understand the functional significance of YTHDF2 interaction with R-loops enriched genomic-loci, m6A and S9.6 DRIP-qPCR assessment of those loci upon YTHDF2 depletion was carried out. Relative to untreated control cells, remarkably, significant enrichment of m6A and S9.6 R-loops were detected in hPSCs with transiently depleted YTHDF2 (figure

8.7.1A-B). As these results suggested a role for YTHDF2 in regulating the stability of RNA:DNA hybrids, the same genomic loci were assessed using S9.6- and m6A-DRIP-qPCR in wild type (WT) and YTHDF2 knockout (KO) HAP1 cells. Consistent with a role for YTHDF2 in regulating the stability of RNA:DNA hybrids, compared to isogenic WT cells, YTHDF2 KO HAP1 cells showed significant accumulation (ranging from 5-50 fold increase) of R-loops in repetitive elements (LINE1s, Alu-S and Alu-Y) and intronic sequences (Figure 8.7.1C-D). In agreement with these results, mass spec analysis of m6A released from RNase H1 treated m6A DRIP experiments show significantly higher m6A levels in YTHDF2 KO HAP1 cells compared to isogenic WT control cells (Figure 8.7.1E-F).

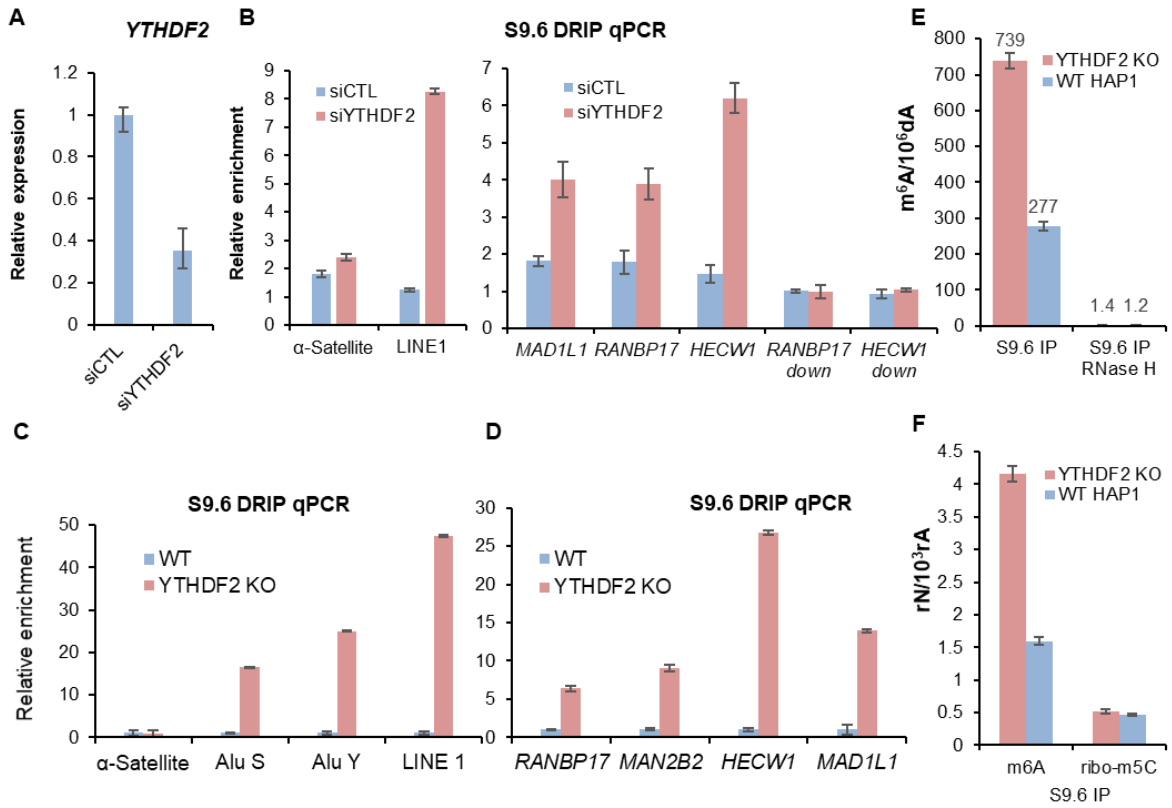



**Figure 8.7.1. YTHDF2 depletion leads to accumulation of R-loops, increased accretion of m6A on RNA:DNA hybrids.** (A) Relative expression of YTHDF2 transcript in hiPSCs transfected with siYTHDF2 siRNAs compared with that in non-targeting control siRNA (siCTL) transfected cells. (B) The results of S9.6 DRIP qPCR of the indicated repeats and intronic sequences performed on siCTL and siYTHDF2 hiPSCs. (C-D) The results of S9.6 DRIP qPCR of the indicated repeats and intronic sequences performed on WT and YTHDF2

KO HAP1 cells. Generic primers amplifying Alu elements from the indicated families and evolutionarily young L1Hs were used. (E-F) SID-UPLC-MS/MS quantification of $m^6A$ in the nucleic acids immuno-precipitated (IP-ed) from *YTHDF2* KO and wild type HAP1 using S9.6 antibody. The results normalized for dA together with S9.6 IP performed on RNase H-pre-treated samples (E) and for ribo-A together with quantification of ribo-$m^5C$ in the IP-ed material (F) are shown. Experimental error is expressed as ± SD.

### 8.7.2. YTHDF2 depletion leads to accumulation of R-loops in human and mouse cells

In agreement with the involvement of YTHDF2 protein in resolving excessive R-loops, YTHDF2 KO cells that express truncated YTHDF2 protein (Figure 8.7.2A) excluded from mitotic chromatin (Figure 8.7.2B) exhibited significantly high S9.6 immunostaining signal, relative to WT cells (Figure 8.7.2C-D). Remarkably, the elevated S9.6 immunostaining signal was consistently detected in mouse *Ythdf2* KO derived neural stem cells (mNSCs) (Figure 8.7.2E-F). Correspondingly, both m6A and S9.6 DRIP qPCR analysis revealed 4-8 fold increase in R-loop enrichment in the open reading frame of mouse LINE1 repeats (Figure 8.7.2G). Taken together, these results demonstrate a conserved role for YTHDF2 in resolving excessive R-loop formation in mammalian cells.

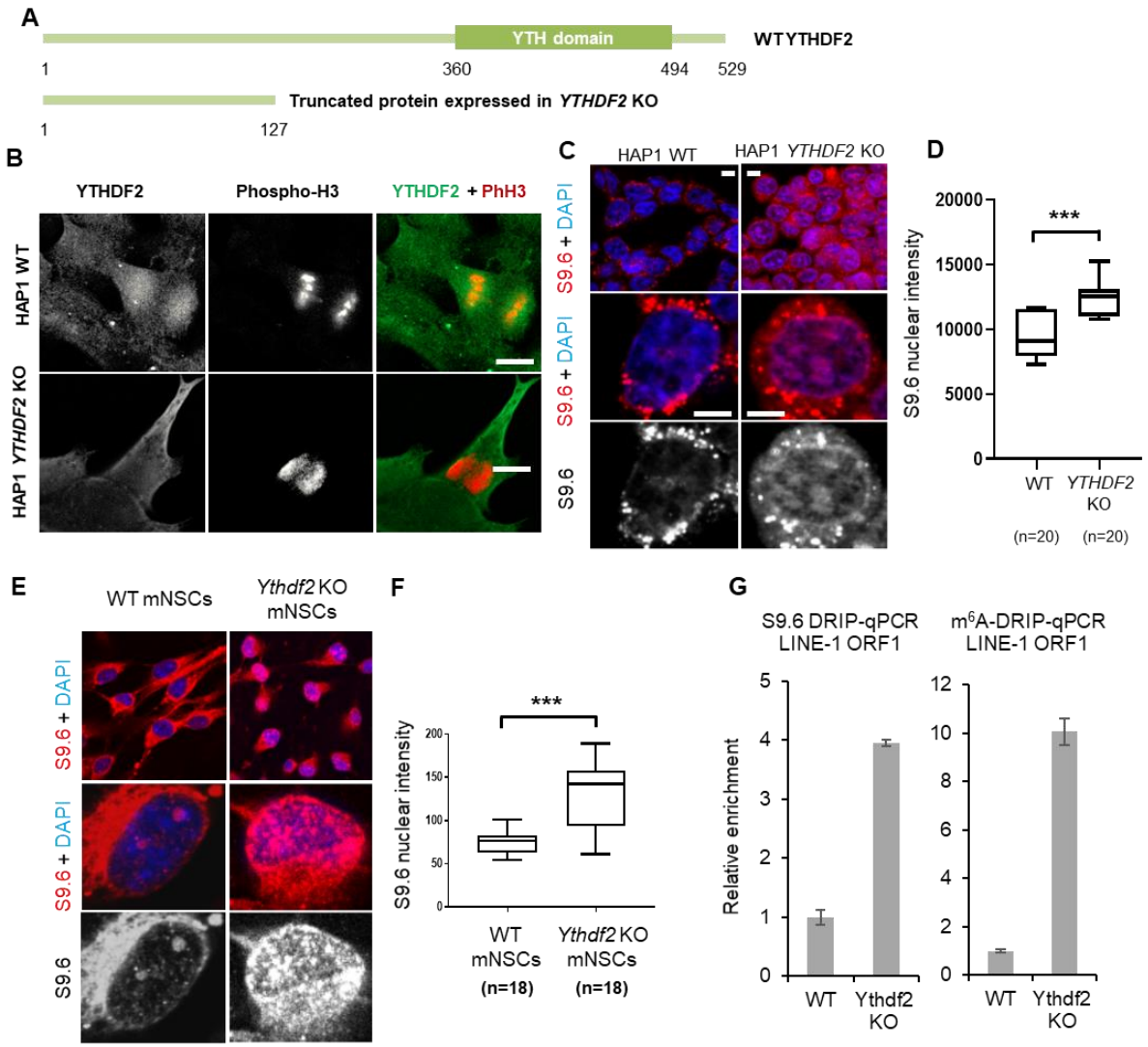



**Figure 8.7.2. YTHDF2 depletion leads to accumulation of R-loops in human and mouse cells**. (A) Schematic illustrating structure of wild type (WT) YTHDF2 protein and truncated protein expressed in *YTHDF2* KO HAP1 cells. The lengths (aa) of the proteins and YTH domain are shown. (B) Representative images of wild type (WT) and *YTHDF2* KO HAP1 cells expressing truncated version of the protein, immunostained for YTHDF2 and phosphorylated histone H3. Individual channels and merged views are shown. (C) Immunostaining of HAP1 wild type and *YTHDF2* KO HAP1 cells with S9.6 antibody and DAPI. Merged images and S9.6 channel views are shown. (D) Quantification of S9.6 signal mean nuclear intensity of cells in (C) The numbers of nuclei used for quantification are indicated in. (E) Immunostaining of wild type/Ythdf2 KO mNSCs with S9.6 antibody and DAPI. (F) Quantification of S9.6 signal mean nuclear intensity of cells in (E). (G) The results of S9.6 and m6A DRIP qPCR of mouse LINE1 ORF1 performed on wild type and Ythdf2 KO mNSCs. Merged images and S9.6 channel views are shown and the numbers of nuclei used for quantification are indicated in (b, f). Representative results of one of 3 independent experiments ± SEM are presented in (c, g), **p < 0.001, ***p < 0.0001.

### 8.8. Summary and conclusion

These results demonstrated the m6A depended interaction of YTHDF2 and HNRNPA2B1 with R-loops. Despite this, depending on the cell-cycle stage, the ability of m6A reader proteins to interact with m6A-modified R-loops significantly differed. Whereas mitotic chromatin seemed devoid of HNRNPA2B1 staining, YTHDF2 exhibited prominent interaction with mitotic chromatin, raising the interesting possibility of whether such interaction is functional. Implying its functional interaction with m6A modified RNA:DNA hybrids, YTHDF2 and S9.6 co-staining signal showed more than 90% overlap across hPSCs nuclei. In agreement with a functional role for YTHDF2 co-localisation with R-loops, both the *in vitro* and *in vivo* analysis showed m6A-dependent YTHDF2 interaction with RNA:DNA hybrids. Interestingly, RNase H or METTL3 depletion mediated increase in R-loop levels led to increased YTHDF2 recruitment to R-loop-containing loci, suggesting functional role for YTHDF2 in R-loop resolution. Corresponding, both transient and constitutive depletion of YTHDF2 resulted in excessive formation of R-loops. Given the novelty of these findings (R-loops accumulation in the absence of YTHDF2), it was very interesting to understand the consequences of YTHDF2 absence on R-loops biology.

# Chapter 9

# Increased DNA damage and cell growth retardation: biological consequences of YTHDF2 depletion.

## 9. Excessive R-loop accumulation, increased DNA damage and cell growth retardation: Biological consequences of YTHDF2 depletion.

The generation of single stranded DNA as part of R-loop formation has been well established to confer vulnerability of the cell to DNA damage (Crossley et al., 2019; García-Muse and Aguilera, 2019) and, as a result, DNA damage has been identified as a characteristic consequence of abnormal R-loop accumulation. This chapter aimed to investigate the significance of YTHDF2 depletion on R-loop metabolism.

### 9.1. YTHDF2 and METTL3 depletion leads to elevated levels of γH2AX phosphorylation in human and mouse cells

Consistent with a role for YTHDF2 in resolving toxic accumulation of R-loops, *YTHDF2* KO HAP1 cells showed a significant increase in immunostaining signal of phosphorylated (at serine 139) histone variant H2AX (γH2AX), a marker for DNA double-strand breaks (Figure 9.1A). The elevated levels of γH2AX staining signal were also detected in the cortex of *Ythdf2* KO embryos and *Ythdf2* KO mNSCs (Figure 9.1C-D). Next, γH2AX ChIP-qPCR analysis was performed to test whether the increased γH2AX staining signal in YTHDF2 KO cells coincides with genomic loci enriched in R-loops. Analysis of γH2AX ChIP-qPCR revealed enrichment of γH2AX in R-loop containing repetitive and intronic sequences (Figure 9.1B). Similarly, siRNA mediated METTL3 depleted hPSCs showed (albeit less significantly) increased γH2AX immunostaining signal (Figure 9.1E). Taken together, these results suggest a novel role for both METTL3 and YTHDF2 in suppressing accumulation of R-loops and ensuing DNA damage.

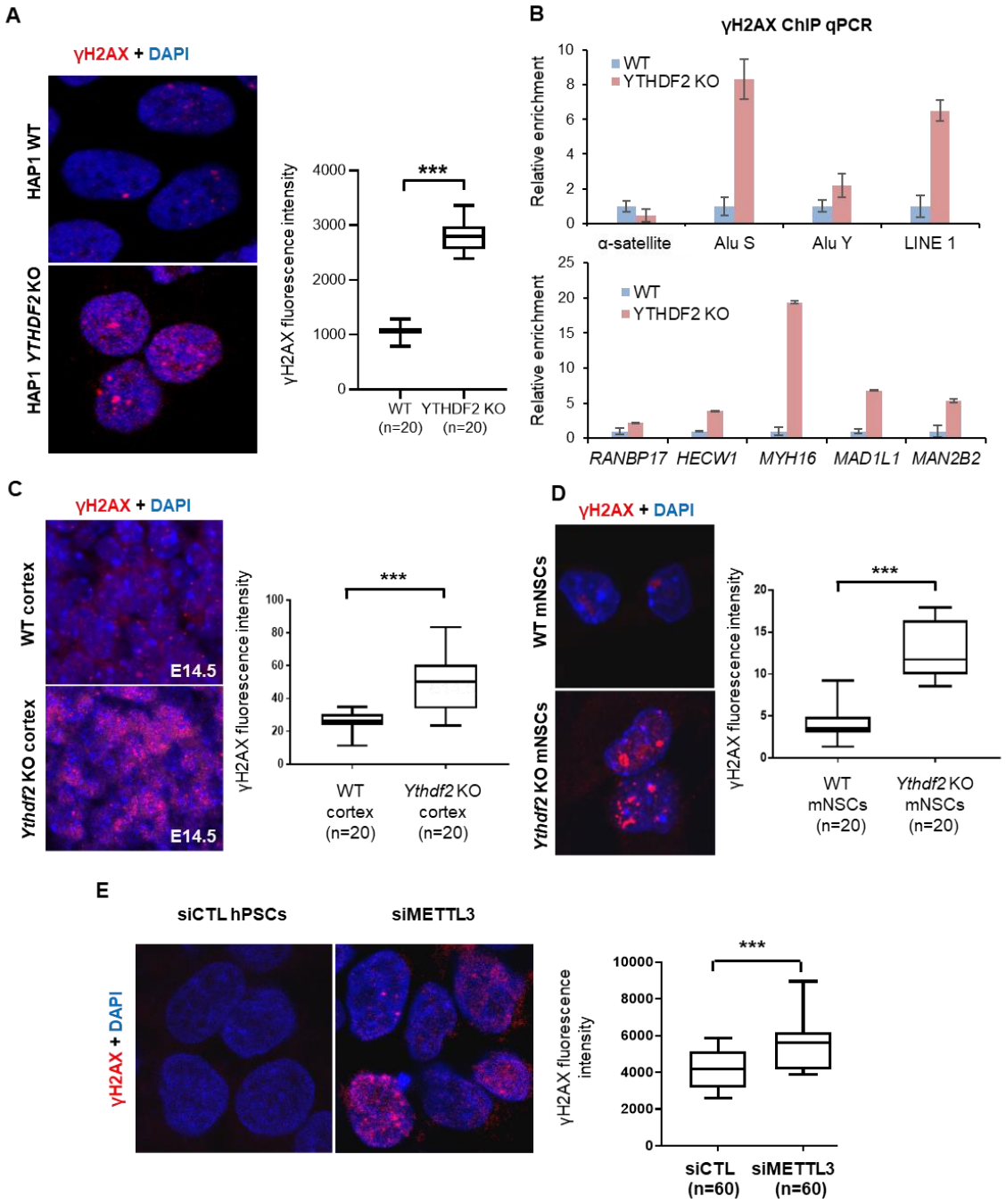



**Figure 9.1. YTHDF2 depletion leads to elevated levels of H2AX phosphorylation in human and mouse cells.** (A) Representative images of *YTHDF2* KO and WT HAP1 cells immunostained for γH2AX and DAPI, and quantification of γH2AX signal intensity in *YTHDF2* KO and WT HAP1 cells. (B) The results of γH2AX ChIP qPCR of the indicated repeats and intronic sequences performed on WT and *YTHDF2* KO HAP1 cells. Generic primers amplifying Alu elements from the indicated families and evolutionarily young L1Hs were used. Data are means of 3 independent experiments ± SEM. (C-D) Representative images of WT/*Ythdf2* KO embryonic brain cortex and mNSCs immunostained for γH2AX and DAPI alongside quantification of γH2AX signal intensity in these tissues/cells. (D) Immunostaining of siCTL and siMETTL hPSCs for γH2AX and DAPI, and quantification of γH2AX signal intensity in these cells *p < 0.05, ***p < 0.0001.

### 9.2. YTHDF-2 depletion lead to cell growth retardation

DNA damage has long been shown to mediate cell growth retardation. Next, the cell growth rate was determined for the DNA damage laden *YTHDF2* KO and isogenic WT HAP1 cells. Only 24 hrs after seeding*YTHDF2* KO cells showed decreased, relative to WT cells, cell number that continued dropping further in the remaining 3 days of culture (Figure 9.2 A). These results are consistent with previous reports of decreased proliferation rate for *Ythdf2* KO-derived neural stem cells (Figure 9.2B) (Li et al., 2018). In sum, these results suggest YTHDF2 involvement in controlling normal cell growth; potentially linked to its role in R-loop resolution.

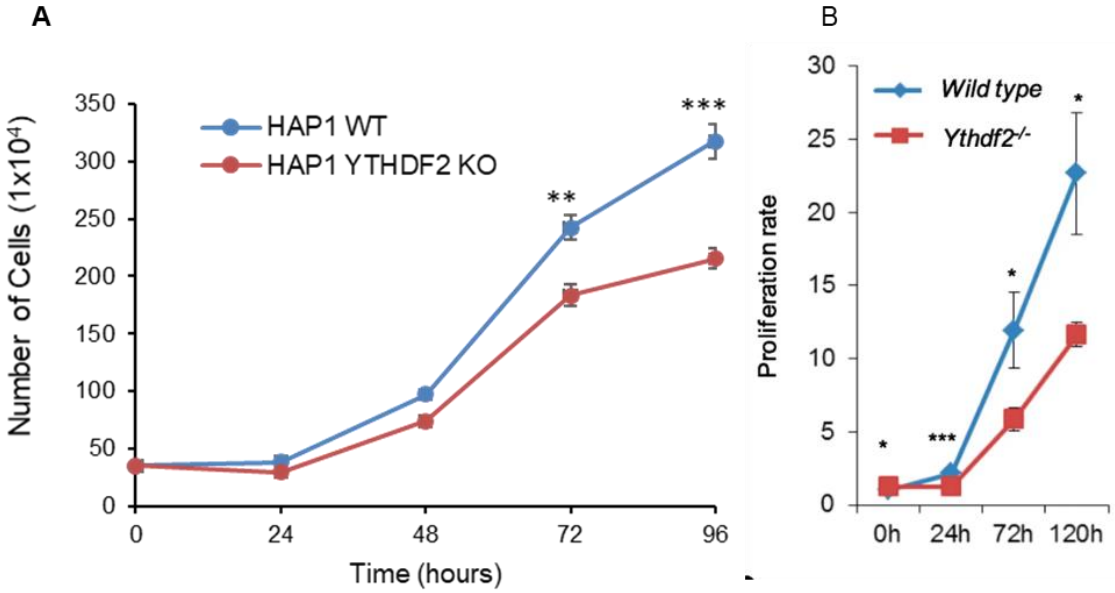



**Figure 9.2. YTHDF2 depletion leads to cell growth retardation in mammalian cells.** (A) The growth curves of WT and *YTHDF2* KO HAP1 cells determined using cell numbers at point of seeding and then every 24hrs until the 4th day. (B) Data from Li *et al.*, 2018 showing Number of viable NSPCs at 0, 24, 72, and 120 h monitored by signal intensity of Presto Blue reagent. Proliferation rate was calculated by normalizing to wild type at 0 h.

### 9.3. RNase H1 overexpression partially rescues increased γH2AX immunostaining signal of YTHDF2 KO cells

A defining feature of R-loops is sensitivity to RNase H treatment (Drolet et al., 1995). By over-expressing nuclear targeted RNase H1 in WT and *YTHDF2* KO, our aim here was to test

whether the increased γH2AX immunostaining signal in *YTHDF2* KO originate from R-loops and thus can be rescued, by over-expression of RNase H. Remarkably, *YTHDF2* KO cells showed significantly reduced γH2AX foci following over-expression of GFP-tagged RNase H1 (Figure 9.3A). In contrast, GFP only transfected cells showed no reduction in γH2AX foci, suggesting reduction of these foci necessitate RNase H over-expression (Figure 9.3B). These results support a conserved function for YTHDF2 in suppressing the build-up of excessive R-loop mediated DNA damage.

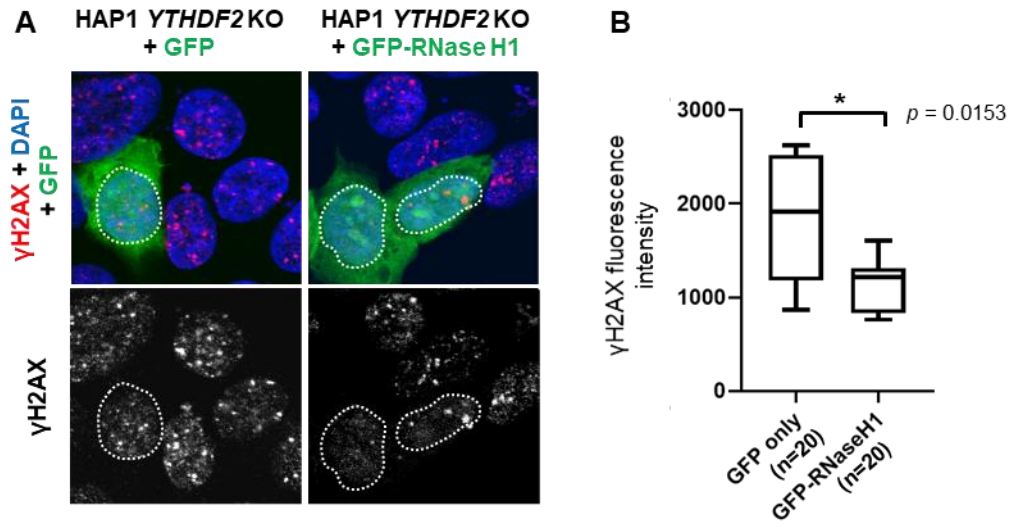



**Figure 9.3. RNase H1 over-expression partially rescues the increased H2AX phosphorylation in YTHDF2 depleted HAP1 cells.** (A) Representative images of *YTHDF2* KO HAP1 cells transfected with GFP-RNase H1 and GFP-only expression constructs immunostained for γH2AX and DAPI, and (B) quantification of γH2AX signal intensity in the GFP-positive cells. P value is indicated. Examples of the nuclei used for signal quantification are marked with dotted shapes. Merged images and γH2AX channel views are shown, and the numbers of nuclei used for quantification are indicated in (B). *p < 0.05.

### 9.4. Summary and conclusion

The striking accumulation of RNA:DNA hybrids in YTHDF2 depleted cells (see chapter 8) raised interesting questions regarding the potential function of YTHDF2 in R-loop biology. Given the growing consensus that accumulation of unscheduled R-loops formation results in

DNA damage (Sollier et al., 2014; Barroso et al., 2019), it raised the intriguing question of whether YTHDF2 contribute to safeguarding genome stability by suppressing R-loop-mediated DNA damage. In support of this, depletion of YTHDF2 led to excessive R-loop formation in mouse and human cells. Importantly, in both human and mouse cells, depletion of YTHDF2 resulted in accumulation of γH2AX foci within R-loop containing genomic loci. To understand whether the accumulated DNA damage in YTHDF2-depleted cells was linked to R-loops, RNase H was over-expressed in YTHDF2 KO human cells. Indicating γH2AX hat the DNA damage in YTHDF2 KO cells is linked to R-loops, RNase H over-expression reduced the accumulation of γH2AX foci. This represented the first described function for YTHDF2 in regulating genome stability via R-loops. Moreover, depletion of YTHDF2 also resulted in cell growth retardation, possibly due to the excessive formation of R-loops and increased rates of DNA damage accumulating in these cells.

# Chapter 10

# General discussion and future directions

**10. General discussions and future directions.**

### 10.1. Discovery of m6A on RNA:DNA hybrids

Motivated by multiple reports of 6mA presence in metazoan genomes (Fu et al., 2015; Greer et al., 2015b; Zhang et al., 2015), initially, this thesis aimed to investigate the presence and potential biological function of DNA 6mA in human cells. Based on the use of two independent approaches, an anti-m6A specific antibody-based highly sensitive immunostaining method and a powerful liquid chromatography-tandem mass spectrometry (LC-MS/MS) capable of detecting as few as 20 modified nucleobases (e.g. 5fC) in the genome (Gackowski et al., 2016; Pfaffeneder et al., 2011), 6mA was not detectable in human cells. These observations are consistent with a growing body of evidence showing that 6mA is either completely absent in mammalian genomes (Schiffers et al., 2017), i.e. occuring beyond the limit of detection of LC-MS/MS (Douvlataniotis et al., 2020; Schiffers et al., 2017), or, according to the more recent studies, may be erroneously incorporated to mammalian DNA by DNA polymerases from bacterial sources (X. Liu et al., 2020). In addition, as 6mA extensively marks prokaryotic genomes, any bacterial (e.g. mycoplasma) contamination may lead to incorrect indication of the presence of physiologically non-negligible levels of 6mA in mammalian genomes (Douvlataniotis et al., 2020; Schiffers et al., 2017) warranting thorough mycoplasma exclusion measures prior to and during 6mA evaluation in nucleic acids from mammalian cells and tissues. Despite maintaining the hPSCs in mycoplasma-free chemically defined culture conditions, consistently, these cells exhibited strong m6A immunostaining signal that was resistant to RNase A (that exclude any contaminating m6A from RNA, (Jia et al., 2011; Meyer et al., 2012; Dominissini et al., 2012; Deng et al., 2015) or DNase I treatments but either significantly decreased or completely lost in cells pre-treated with RNase H, a ribo-endonuclease that specifically targets the RNA component of RNA:DNA hybrids for degradation (Hyjek et al., 2019) . Moreover, using dot-blot experiments, more recently, Yang and colleagues also confirmed the presence of m6A on RNA:DNA (X. Yang et al., 2019)

although no quantitative assessment was shown. Despite re-confirming the specificity of the anti-m6A antibody in this thesis and by other labs (Mondo et al., 2017; Zhao et al., 2017), it is conceivable that the anti-m6A antibody could also exhibit binding affinity, albeit weakly, to other structurally similar RNA modifications (Linder et al., 2015). Therefore, in collaboration with the Gackowski lab in Poland, an antibody-independent, quantitative assessment of RNase H of m6A in hPSCs DNA was performed, which independently demonstrated the presence of m6A marked RNA:DNA hybrids. In sum, these results provide the first observation of a noncanonical nucleic base in R-loops, that have recently been independently confirmed in a number of studies (X. Yang et al., 2019; Zhang et al., 2020). Given the wide-ranging regulatory roles m6A is known to play in RNA biology (Fu et al., 2014) coupled with the conserved regulatory function of R-loops (Niehrs and Luke, 2020), it is anticipated that the discovery of m6A presence on R-loops will likely impact all steps of R-loop/RNA:DNA hybrid biogenesis and possibly extend the biological function of m6A itself, thus providing an insight into another layer of gene regulation. Moreover, the discovery of m6A on R-loops might be particularly helpful in settling the ongoing debate around the highly contentious 6mA presence in mammalian genomes (Koziol et al., 2016; Wu et al., 2016a; Liu et al., 2016; Mendel et al., 2018)(Koziol et al., 2016; Liu et al., 2016; Wu et al., 2016b; Schiffers et al., 2017; Yao et al., 2017; Stöger, 2017; Douvlataniotis et al., 2020). Thus, the discovery of m6A on R-loops suggest that re-evaluation of all previously reported data on the 6mA presence in mammalian genomes is necessary. Notably, to date, all reports of 6mA detection in mammalian genomes are presented without RNase H control and, as the result, m6A mapping on R-loops might be interpreted as 6mA (Wu et al., 2016b). Moreover, as the anti-m6A antibody can also bind DNA 6mA and, unless genomic DNA is harvested from contamination-free environment, the presence of 6mA in contaminating species e.g. bacteria could confound interpretation of 6mA analysis in mammalian genomes. Furthermore, the stability of RNA:DNA hybrids is higher

than that of double stranded DNA (Roberts and Crothers, 1992), and given the resistance of RNA:DNA hybrids to RNase A treatment (Allison and Wang, 2019), any results of 6mA immunostaining in the absence of RNase H may erroneously indicate the presence of 6mA.

### 10.2. Genomic distribution of m6A marked RNA:DNA hybrids.

As soon as the presence of m6A on RNA:DNA hybrids has been confirmed, the aim was to obtain insights into the biological function of m6A on R-loops. To this end, by adapting a previously published genome-wide RNA:DNA hybrid mapping protocol (Ginno et al., 2012), the genome-wide distribution of m6A-marked RNA:DNA hybrids in hPSCs was obtained where the corresponding sequences showed RNase H sensitivity in control samples. Interestingly, the global map of m6A marked RNA:DNA hybrids showed almost identical profile to the R-loops maps obtained with the S9.6 antibody, suggesting that m6A marks most of RNA:DNA hybrids. Although this is somehow consistent with previous finding showing the presence of m6A on the majority of messenger RNAs (Bokar, 2005; Linder et al., 2015), the anti-m6A antibody captured 3 folds more RNA:DNA compared with S9.6 antibody specific for RNA:DNA hybrids. Using synthetic m6A modified and unmodified RNA:DNA hybrids as spike-in control during m6A and S9.6 immunoprecipitation reactions, the affinity of m6A antibody was shown to be considerably higher (3 – 5 folds) than S9.6 for the same hybrid. This is consistent with previous reports that show that various antibodies raised to the same target may have considerable differences in their affinities (Shim, 2011), a widely recognised draw-back of therapeutic application of many antibodies. Nevertheless, the recent development of an antibody-independent approach for mapping genome-wide distribution of R-loops can provide a potential alternative for studying the genomic distribution of RNA:DNA hybrids (Malig et al., 2020), however, in the context of m6A marked R-loops, such an approach may require coupling with an antibody-independent mapping of m6A-marked RNA:DNA hybrids.

Consistent with co-transcriptional formation of RNA:DNA hybrids (Huertas and Aguilera, 2003; Stork et al., 2016), the genome-wide profiling of m6A-modified R-loops revealed a strong enrichment of m6A marked RNA:DNA hybrids in transcribed regions. Strikingly, nearly 50% of m6A- and S9.6- captured RNA:DNA hybrids were mapped to transcribed intergenic regions which are increasingly being recognised to play important and often conserved regulatory functions (Pheasant and Mattick, 2007; Hainer et al., 2011; Mattick, 2018). Importantly, these data are consistent with several maps of R-loops in various model systems (Nadel et al., 2015; Wahba et al., 2016; Al-Hadid and Yang, 2016; Chen et al., 2017). In addition to intergenic regions, more than 20% of mapped m6A marked RNA:DNA hybrids were enriched in intronic sequences, suggesting potential involvement of these R-loops in regulating RNA processing (Li and Manley, 2005). In agreement with these results, recent analysis implicated intronic R-loops in regulating alternative RNA splicing (El Hage et al., 2014; Wan et al., 2015; Nguyen et al., 2018; Chen et al., 2018) and genome instability as a consequence of their accumulation (Chan et al., 2014; Nguyen et al., 2019; Tam and Stirling, 2019). Moreover, a substantial number of m6A and S9.6 peaks mapped to either promoter, 5'UTR or 3'UTR regions, suggesting direct involvement in transcriptional regulation (Niehrs and Luke, 2020), and this is consistent with previous studies demonstrating R-loop enrichment in these regions (Ginno et al., 2012; Skourti-Stathaki et al., 2014; Sanz et al., 2016a; Dumelie and Jaffrey, 2017). In support of this, promoter-associated R-loops have been shown to promote demethylation by hindering access of DNA methyltransferases to their target CpG sequences (Arab et al., 2019) or altering chromatin structure to either promote antisense transcription (Tan-Wong et al., 2019) or induce repressive chromatin structure (Sun et al., 2013; Colak et al., 2014). Moreover, formation of R-loops at 3'UTR pause sites have been associated with transcriptional termination (Gromak et al., 2006; Skourti-Stathaki et al., 2011). Unlike some of the unusually long (up to 1kb) R-loops reported in some studies (Sanz et al.,

2016b), the majority of the RNA:DNA hybrids obtained in this study are less than 500 bp in length. This is consistent with the fact that, apart from R-loops reported in B-cells formed during class switch recombination (Yu et al., 2003), the majority of R-loops ranges from 80 to 500 bp according to a number of reports (Sun et al., 2013; El Hage et al., 2014; Skourti-Stathaki and Proudfoot, 2014; Costantino and Koshland, 2015; Wahba et al., 2016; García-Muse and Aguilera, 2019). Notwithstanding the differences in the model system used for R-loop mapping, technical variation in R-loop mapping, particularly in the fragmentation step of DNA prior to immunoprecipitation could contribute to differences in the length of reported R-loops (Sanz et al., 2016b; Wahba et al., 2016). Consistent with this, it was found that the use of restriction-enzyme based DNA fragmentation could result in over-estimation of R-loop size as they are known to predominantly target the repeat-rich non-coding sequences (Halász et al., 2017). Nevertheless, the recent development of high-resolution R-loop mapping techniques might aid the standardization of methods used for R-loop mapping (Dumelie and Jaffrey, 2017; Malig et al., 2020). Moreover, the m6A RNA:DNA hybrid mapping approach described in this thesis provide a robust and highly sensitive method for profiling R-loops genome-wide or at the level of a single-locus.

### 10.3. Lack of correlation between m6A marked RNA:DNA hybrid and major histone marks.

To gain insights into the chromatin environment associated with m6A-marked RNA:DNA hybrids, both m6A and S9.6 derived RNA:DNA hybrids were overlapped with previously published histone marks in hPSCs. Despite previous indications of particular histone modifications and R-loop formation (Skourti-Stathaki et al., 2014; Zeller et al., 2016; Sanz et al., 2016a), there was not obvious correlation between either m6A or S9.6 derived RNA:DNA hybrids. Consistent with this finding, the presence of RNA:DNA hybrids has previously been shown to disfavour the presence of histone marks (Dunn and Griffith, 1980; Crossley et al.,

2019). Moreover, recent analysis suggests that the association of RNA:DNA hybrids with particular histone marks is more indicative of non-physiological R-loop formation (García-Pichardo et al., 2017). In summary, these results suggest a model whereby m6A marked RNA:DNA hybrids provide a separate yet previously unknown chromatin environment that may further alter both the local chromatin topology and confirmation, to either exclude or promote specific chromatin interacting factors.

### 10.4. The cell-cycle as potential regulator of R-loop dynamics

Compared to interphase cells, mitotic cells consistently exhibited RNase A resistant high m6A immunostaining signal, and in an aim to probe this further, hPSCs were sorted using flow cytometry into G0/G1, S and G2/M cell cycle populations based on propidium iodide staining intensity, before assessment of m6A modified RNA:DNA hybrid levels. Interestingly, accumulation of RNA:DNA hybrids varied depending on the stage of the cell cycle for both intronic and retroelements sequences. Consistently with the immunostaining results, both m6A and S9.6 derived RNA:DNA hybrids were enriched in S and G2/M phases of the cell cycle with noticeable depletion in interphase cells. While confirming previous reports showing variation in R-loop levels between dividing and nondividing cells (Gómez-González et al., 2009; Graf et al., 2017; Kabeche et al., 2018), these results also extend those findings by showing mitotic enrichment of R-loops in LINE1 retro-elements, potentially regulating their retroposition in dividing cells (Shi et al., 2007; Xie et al., 2013). As waves of gene reactivation is known to occur during and on-exit from mitosis, it would be particularly interesting to further dissect the role of m6A marked RNA:DNA hybrids in LINE1 expression (Palozola et al., 2017), especially in the context of pluripotency, and cell fate decisions, wherein LINE1 mobilization has recently been implicated (Klawitter et al., 2016; Percharde et al., 2018; Flasch et al., 2019; Wang et al., 2020).

### 10.5. Modulation of m6A levels uncover its role in regulating R-loop stability

Technical difficulties of selectively modifying m6A has previously challenged attempts into understanding the precise biological function of m6A (Linder et al., 2015). The discovery of m6A writers, readers and eraser proteins not only demonstrated its dynamic nature (Fu et al., 2014) but also provided an excellent tool to probe the function of m6A. Consequently, m6A has been implicated in almost all aspects of RNA metabolism at the cellular level from pluripotency and differentiation, to development and disease (Bokar, 2005; Dominissini et al., 2012; Fu et al., 2014; Geula et al., 2015; Xu et al., 2017; Barbieri and Kouzarides, 2020; Begik et al., 2020). One of the most striking consequences of m6A deposition on RNAs (Desrosiers et al., 1974; Perry and Kelley, 1975) is its ability to reduce their half-life (Sommer et al., 1978; Ke et al., 2017) and catalyse their eventual destabilization (Wang et al., 2014; Zaccara and Jaffrey, 2020; Lasman et al., 2020; Lee et al., 2020). The finding that m6A marks the majority of RNA:DNA hybrids (see chapter 4) and demonstration that RNA:DNA hybrids exhibit cell cycle dependent dynamics (see chapter 6), raised the interesting hypothesis of whether m6A, similar to the consequence of its presence on mRNAs (Wang et al., 2014), also regulates the stability of RNA:DNA hybrids. Consistent with previous analysis, transient depletion of METTL3 in hPSCs led to reduced levels (29%) of m6A deposition as assessed by mass spectrometry (Liu et al., 2014). While such reduction of m6A deposition did not seem to significantly affect expression of R-loop containing loci (see chapter 7), remarkably, it led to 2-6 fold increase in the level of R-loops detected in intronic sequences and LINE1 retroelements during S and G2/M phases of the cell cycle compared to control hPSCs. These results not only demonstrate the involvement of m6A in regulating the stability of RNA:DNA hybrids but are also consistent with other studies implicating m6A in promoting rapid turnover of target RNAs (Liu et al., 2014; Wang et al., 2014; Ke et al., 2017; Lee et al., 2020). In accordance with this, m6A-decorated RNA:DNA hybrids have recently been implicated in

promoting homologous recombination-mediated double strand break repair (Zhang et al., 2020). Moreover, the finding that nuclear m6A regulate RNA:DNA hybrid stability present a novel nuclear function for m6A beyond its much-studied role in regulating cytoplasmic RNA. In agreement with this, recently, the He lab showed that modulation of m6A could also impact not only the distribution but also levels of chromatin-associated RNAs (J. Liu et al., 2020). In addition, m6A could contribute to reduction of the thermodynamic stability of RNA:DNA hybrids (Roberts and Crothers, 1992) that is defined by unique structure of their double helix intermediate between A-form characteristic for RNA:RNA duplexes and B-form usual for double stranded DNA (Shaw and Arya, 2008).

To gain further insights into the biological significance of m6A on R-loops, the best characterised m6A reader proteins were probed for their interaction with m6A modified R-loops. This includes, YTHDF1 (Wang et al., 2015), YTHDF2 (Wang et al., 2014) and HNRNPA2B1 (Alarcón et al., 2015) that exert their function on RNA in m6A dependent manner along with METTL3 (Liu et al., 2014) were found to physically interact with native R-loops. In support of this, these proteins have already been reported in previous studies of RNA:DNA hybrid protein interactomes (Cristini et al., 2018; Wang et al., 2018). As m6A modified RNA:DNA hybrids showed preferential enrichment in mitotic cells, the spatial distribution of these reader proteins was examined using immunostaining experiments (see chapter 8). Although all 3 key m6A reader proteins were detected in interphase hPSCs nuclei, only YTHDF2 showed strong localisation to mitotic chromatin, suggesting its functional role during mitosis. In support of this, YTHDF2 was earlier found amongst proteins interacting with cell-cycle resolved mitotic chromatin (Ginno et al., 2018), however no functional analysis of its role in the chromatin was presented. To investigate whether YTHDF2 translocation to mitotic chromatin is dependent on m6A, we assessed YTHDF2 binding affinity to synthetic m6A-containing and non-modified RNA:DNA hybrids using micro-scale thermophoresis and

EMSA. In both these experiments, YTHDF2 exhibited stronger affinity to m6A-modified synthetic RNA:DNA hybrids compared to unmodified hybrids. In agreement with this, YTHDF2 exhibited strong co-localization with S9.6 in immunostaining experiments, further suggesting a functional interaction between YTHDF2 and R-loops. Consistent with this, YTHDF2 chromatin immunoprecipitation analysis showed recruitment of YTHDF2 to R-loops containing genomic loci. In order to probe the function of YTHDF2 recruitment to sites of R-loops, next, profiles of RNA:DNA hybrids in intronic and LINE1 retroelement regions were assessed following depletion of YTHDF2. In these experiments, YHTDF2 depletion led to increased accumulation of m6A modified RNA:DNA hybrids, suggesting the involvement of YTHDF2 in resolution of RNA:DNA hybrids. Consistent with this, depletion of both YTHDF2 and METTL3 led to increase in γH2AX immunostaining signal, a marker of double strand DNA breaks (Georgoulis et al., 2017) and a characteristic feature of persistent R-loops (Hamperl et al., 2017; Cristini et al., 2019; Crossley et al., 2019). Taken together, the data presented in this thesis support a role for m6A in regulating the stability of RNA:DNA hybrids that accumulate during specific stages of the cell cycle. In this context, m6A through its reader proteins might regulate YTHDF2 that might require the recruitment of additional factors such as helicases (Cristini et al., 2018; Wang et al., 2018) or endonucleases (Park et al., 2019) to achieve R-loop resolution. As a result, YTHDF2 could help ensure proper transition through the cell cycle (Fei et al., 2020). Consistent with this, YTHDF2 was found among genes whose down-regulation led to significant increase in γH2AX in human cells (Paulsen et al., 2009). The thesis envisages future implication of m6A in virtually all instances of R-loop formation, scanning, and resolution ranging from physiological to pathological contexts. Moreover, the presence of m6A-marked R-loops in human cells will likely contribute to understanding the exact molecular mechanisms of balancing the paradoxical roles of R-loops between physiological phenomena and pathological contexts that threaten genome integrity.

Furthermore, ultimately, recent advances in RNA editing (Liu et al., 2019) might allow using m6A and other RNA modifications for precise programming of the stability of RNA:DNA hybrids that may be instrumental for their therapeutic manipulation in R-loop mediated pathologies (Richard and Manley, 2017).

## References:


Aas, P.A., Otterlei, M., Falnes, P.O., Vågbø, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjørås, M., Slupphaug, G., Seeberg, E., Krokan, H.E., 2003. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421, 859–863. https://doi.org/10.1038/nature01363

Abakir, A., Wheldon, L., Johnson, A.D., Laurent, P., Ruzov, A., 2016. Detection of Modified Forms of Cytosine Using Sensitive Immunohistochemistry. J. Vis. Exp. JoVE. https://doi.org/10.3791/54416

Achwal, C.W., Iyer, C.A., Chandra, H.S., 1983. Immunochemical evidence for the presence of 5mC, 6mA and 7mG in human, Drosophila and mealybug DNA. FEBS Lett. 158, 353–358. https://doi.org/10.1016/0014-5793(83)80612-7

Agarwala, S.D., Blitzblau, H.G., Hochwagen, A., Fink, G.R., 2012. RNA methylation by the MIS complex regulates a cell fate decision in yeast. PLoS Genet. 8, e1002732. https://doi.org/10.1371/journal.pgen.1002732

Aguilera, A., García-Muse, T., 2012. R loops: from transcription byproducts to threats to genome stability. Mol. Cell 46, 115–124. https://doi.org/10.1016/j.molcel.2012.04.009

Aguilera, A., Gómez-González, B., 2008. Genome instability: a mechanistic view of its causes and consequences. Nat. Rev. Genet. 9, 204–217. https://doi.org/10.1038/nrg2268

Alarcón, C.R., Lee, H., Goodarzi, H., Halberg, N., Tavazoie, S.F., 2015. N6-methyladenosine marks primary microRNAs for processing. Nature 519, 482–485. https://doi.org/10.1038/nature14281

Alecki, C., Chiwara, V., Sanz, L.A., Grau, D., Arias Pérez, O., Boulier, E.L., Armache, K.-J., Chédin, F., Francis, N.J., 2020. RNA-DNA strand exchange by the Drosophila Polycomb complex PRC2. Nat. Commun. 11, 1781. https://doi.org/10.1038/s41467-020-15609-x

Al-Hadid, Q., Yang, Y., 2016. R-loop: an emerging regulator of chromatin dynamics. Acta Biochim. Biophys. Sin. 48, 623–631. https://doi.org/10.1093/abbs/gmw052

Allison, D.F., Wang, G.G., 2019. R-loops: formation, function, and relevance to cell stress. Cell Stress 3, 38–46. https://doi.org/10.15698/cst2019.02.175

Ammermann, D., Steinbrück, G., Baur, R., Wohlert, H., 1981. Methylated bases in the DNA of the ciliate Stylonychia mytilus. Eur. J. Cell Biol. 24, 154–156.

Arab, K., Karaulanov, E., Musheev, M., Trnka, P., Schäfer, A., Grummt, I., Niehrs, C., 2019. GADD45A binds R-loops and recruits TET1 to CpG island promoters. Nat. Genet. 51, 217–223. https://doi.org/10.1038/s41588-018-0306-6

Arber, W., Dussoix, D., 1962. Host specificity of DNA produced by Escherichia coli. I. Host controlled modification of bacteriophage lambda. J. Mol. Biol. 5, 18–36. https://doi.org/10.1016/s0022-2836(62)80058-8

Avery, O.T., Macleod, C.M., McCarty, M., 1944. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. J. Exp. Med. 79, 137–158. https://doi.org/10.1084/jem.79.2.137

Baaklini, I., Usongo, V., Nolent, F., Sanscartier, P., Hraiky, C., Drlica, K., Drolet, M., 2008. Hypernegative supercoiling inhibits growth by causing RNA degradation. J. Bacteriol. 190, 7346–7356. https://doi.org/10.1128/JB.00680-08

Bader, A.S., Bushell, M., 2020. DNA:RNA hybrids form at DNA double-strand breaks in transcriptionally active loci. Cell Death Dis. 11, 280. https://doi.org/10.1038/s41419-020-2464-6

Baldacci, G., Chérif-Zahar, B., Bernardi, G., 1984. The initiation of DNA replication in the mitochondrial genome of yeast. EMBO J. 3, 2115–2120.

Bao, W., Kojima, K.K., Kohany, O., 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA 6. https://doi.org/10.1186/s13100-015-0041-9
Barbieri, I., Kouzarides, T., 2020. Role of RNA modifications in cancer. Nat. Rev. Cancer 20, 303–322. https://doi.org/10.1038/s41568-020-0253-2
Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S.C., Aspris, D., Migliori, V., Bannister, A.J., Han, N., De Braekeleer, E., Ponstingl, H., Hendrick, A., Vakoc, C.R., Vassiliou, G.S., Kouzarides, T., 2017. Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. Nature 552, 126–131. https://doi.org/10.1038/nature24678
Barroso, S., Herrera-Moyano, E., Muñoz, S., García-Rubio, M., Gómez-González, B., Aguilera, A., 2019. The DNA damage response acts as a safeguard against harmful DNA-RNA hybrids of different origins. EMBO Rep. 20, e47250. https://doi.org/10.15252/embr.201847250
Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., Carter, A.C., Flynn, R.A., Zhou, C., Lim, K.-S., Dedon, P., Wernig, M., Mullen, A.C., Xing, Y., Giallourakis, C.C., Chang, H.Y., 2014. m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15, 707–719. https://doi.org/10.1016/j.stem.2014.09.019
Batzer, M.A., Deininger, P.L., 2002. Alu repeats and human genomic diversity. Nat. Rev. Genet. 3, 370–379. https://doi.org/10.1038/nrg798
Bayona-Feliu, A., Casas-Lamesa, A., Reina, O., Bernués, J., Azorín, F., 2017. Linker histone H1 prevents R-loop accumulation and genome instability in heterochromatin. Nat. Commun. 8, 283. https://doi.org/10.1038/s41467-017-00338-5
Begik, O., Lucas, M.C., Liu, H., Ramirez, J.M., Mattick, J.S., Novoa, E.M., 2020. Integrative analyses of the RNA modification machinery reveal tissue- and cancer-specific signatures. Genome Biol. 21, 97. https://doi.org/10.1186/s13059-020-02009-z
Beh, L.Y., Debelouchina, G.T., Clay, D.M., Thompson, R.E., Lindblad, K.A., Hutton, E.R., Bracht, J.R., Sebra, R.P., Muir, T.W., Landweber, L.F., 2019. Identification of a DNA N6-Adenine Methyltransferase Complex and Its Impact on Chromatin Organization. Cell 177, 1781-1796.e25. https://doi.org/10.1016/j.cell.2019.04.028
Belotserkovskii, B.P., Soo Shin, J.H., Hanawalt, P.C., 2017. Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter. Nucleic Acids Res. 45, 6589–6599. https://doi.org/10.1093/nar/gkx403
Bertero, A., Brown, S., Madrigal, P., Osnato, A., Ortmann, D., Yiangou, L., Kadiwala, J., Hubner, N.C., de Los Mozos, I.R., Sadée, C., Lenaerts, A.-S., Nakanoh, S., Grandy, R., Farnell, E., Ule, J., Stunnenberg, H.G., Mendjan, S., Vallier, L., 2018. The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. Nature 555, 256–259. https://doi.org/10.1038/nature25784
Bhowmik, S., Krishnamurthy, R., 2019. The role of sugar-backbone heterogeneity and chimeras in the simultaneous emergence of RNA and DNA. Nat. Chem. 11, 1009–1018. https://doi.org/10.1038/s41557-019-0322-x
Björk, G.R., Ericson, J.U., Gustafsson, C.E., Hagervall, T.G., Jönsson, Y.H., Wikström, P.M., 1987. Transfer RNA modification. Annu. Rev. Biochem. 56, 263–287. https://doi.org/10.1146/annurev.bi.56.070187.001403
Blow, M.J., Clark, T.A., Daum, C.G., Deutschbauer, A.M., Fomenkov, A., Fries, R., Froula, J., Kang, D.D., Malmstrom, R.R., Morgan, R.D., Posfai, J., Singh, K., Visel, A., Wetmore, K., Zhao, Z., Rubin, E.M., Korlach, J., Pennacchio, L.A., Roberts, R.J., 2016. The Epigenomic Landscape of Prokaryotes. PLoS Genet. 12, e1005854. https://doi.org/10.1371/journal.pgen.1005854
Boguslawski, S.J., Smith, D.E., Michalak, M.A., Mickelson, K.E., Yehle, C.O., Patterson, W.L., Carrico, R.J., 1986. Characterization of monoclonal antibody to DNA.RNA and its application to

immunodetection of hybrids. J. Immunol. Methods 89, 123–130. https://doi.org/10.1016/0022-1759(86)90040-2
Bokar, J., 2005. The biosynthesis and functional roles of methylated nucleosides in eukaryotic mRNA, in: Fine-Tuning of RNA Functions by Modification and Editing. pp. 141–177. https://doi.org/10.1007/b106365
Bokar, J.A., Rath-Shambaugh, M.E., Ludwiczak, R., Narayan, P., Rottman, F., 1994. Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. J. Biol. Chem. 269, 17697–17704.
Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., Rottman, F.M., 1997. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA N. Y. N 3, 1233–1247.
Boque-Sastre, R., Soler, M., Guil, S., 2017. Detection and Characterization of R Loop Structures. Methods Mol. Biol. Clifton NJ 1543, 231–242. https://doi.org/10.1007/978-1-4939-6716-2_13
Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S., Villanueva, A., Esteller, M., Guil, S., 2015. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. Proc. Natl. Acad. Sci. U. S. A. 112, 5785–5790. https://doi.org/10.1073/pnas.1421197112
Boulé, J.-B., Zakian, V.A., 2007. The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. Nucleic Acids Res. 35, 5809–5818. https://doi.org/10.1093/nar/gkm613
Boye, E., Løbner-Olesen, A., Skarstad, K., 2000. Limiting DNA replication to once and only once. EMBO Rep. 1, 479–483. https://doi.org/10.1093/embo-reports/kvd116
Camacho, E.M., Serna, A., Madrid, C., Marqués, S., Fernández, R., de la Cruz, F., Juárez, A., Casadesús, J., 2005. Regulation of finP transcription by DNA adenine methylation in the virulence plasmid of Salmonella enterica. J. Bacteriol. 187, 5691–5699. https://doi.org/10.1128/JB.187.16.5691-5699.2005
Cantara, W.A., Crain, P.F., Rozenski, J., McCloskey, J.A., Harris, K.A., Zhang, X., Vendeix, F.A.P., Fabris, D., Agris, P.F., 2011. The RNA Modification Database, RNAMDB: 2011 update. Nucleic Acids Res. 39, D195-201. https://doi.org/10.1093/nar/gkq1028
Carles-Kinch, K., Kreuzer, K.N., 1997. RNA-DNA hybrid formation at a bacteriophage T4 replication origin. J. Mol. Biol. 266, 915–926. https://doi.org/10.1006/jmbi.1996.0844
Carroll, S.M., Narayan, P., Rottman, F.M., 1990. N6-methyladenosine residues in an intron-specific region of prolactin pre-mRNA. Mol. Cell. Biol. 10, 4456–4465. https://doi.org/10.1128/mcb.10.9.4456
Carroll, T.S., Liang, Z., Salama, R., Stark, R., de Santiago, I., 2014. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. Front. Genet. 5, 75. https://doi.org/10.3389/fgene.2014.00075
Castellano-Pozo, M., García-Muse, T., Aguilera, A., 2012. R-loops cause replication impairment and genome instability during meiosis. EMBO Rep. 13, 923–929. https://doi.org/10.1038/embor.2012.119
Catalan, N., Selz, F., Imai, K., Revy, P., Fischer, A., Durandy, A., 2003. The block in immunoglobulin class switch recombination caused by activation-induced cytidine deaminase deficiency occurs prior to the generation of DNA double strand breaks in switch mu region. J. Immunol. Baltim. Md 1950 171, 2504–2509. https://doi.org/10.4049/jimmunol.171.5.2504
Cerritelli, S.M., Crouch, R.J., 2009. Ribonuclease H: the enzymes in eukaryotes. FEBS J. 276, 1494–1505. https://doi.org/10.1111/j.1742-4658.2009.06908.x
Chan, Y.A., Aristizabal, M.J., Lu, P.Y.T., Luo, Z., Hamza, A., Kobor, M.S., Stirling, P.C., Hieter, P., 2014. Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip. PLoS Genet. 10. https://doi.org/10.1371/journal.pgen.1004288

Chang, M., Lv, H., Zhang, W., Ma, C., He, X., Zhao, S., Zhang, Z.-W., Zeng, Y.-X., Song, S., Niu, Y., Tong, W.-M., 2017. Region-specific RNA m6A methylation represents a new layer of control in the gene regulatory network in the mouse brain. Open Biol. 7. https://doi.org/10.1098/rsob.170166

Chang, V.Y., Federman, N., Martinez-Agosto, J., Tatishchev, S.F., Nelson, S.F., 2013. Whole exome sequencing of pediatric gastric adenocarcinoma reveals an atypical presentation of Li-Fraumeni syndrome. Pediatr. Blood Cancer 60, 570–574. https://doi.org/10.1002/pbc.24316

Chaudhuri, J., Alt, F.W., 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat. Rev. Immunol. 4, 541–552. https://doi.org/10.1038/nri1395

Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E., Alt, F.W., 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature 422, 726–730. https://doi.org/10.1038/nature01574

Chávez, S., Beilharz, T., Rondón, A.G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J.Q., Lithgow, T., Aguilera, A., 2000. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. EMBO J. 19, 5824–5834. https://doi.org/10.1093/emboj/19.21.5824

Chen, K., Zhao, B.S., He, C., 2016. Nucleic Acid Modifications in Regulation of Gene Expression. Cell Chem. Biol. 23, 74–85. https://doi.org/10.1016/j.chembiol.2015.11.007

Chen, L., Chen, J.-Y., Huang, Y.-J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., He, S., Zhou, Y., Abdel-Wahab, O., Zhang, D.-E., Fu, X.-D., 2018. The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. Mol. Cell 69, 412-425.e6. https://doi.org/10.1016/j.molcel.2017.12.029

Chen, L., Chen, J.-Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H., Zhou, Y., Zhang, D.-E., Fu, X.-D., 2017. R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. Mol. Cell 68, 745-757.e5. https://doi.org/10.1016/j.molcel.2017.10.008

Chen, S., Zhou, L., Wang, Y., 2020. ALKBH5-mediated m6A demethylation of lncRNA PVT1 plays an oncogenic role in osteosarcoma. Cancer Cell Int. 20, 34. https://doi.org/10.1186/s12935-020-1105-6

Cheng, S.C., Herman, G., Modrich, P., 1985. Extent of equilibrium perturbation of the DNA helix upon enzymatic methylation of adenine residues. J. Biol. Chem. 260, 191–194.

Clancy, M.J., Shambaugh, M.E., Timpte, C.S., Bokar, J.A., 2002. Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. Nucleic Acids Res. 30, 4509–4518. https://doi.org/10.1093/nar/gkf573

Cloutier, S.C., Wang, S., Ma, W.K., Al Husini, N., Dhoondia, Z., Ansari, A., Pascuzzi, P.E., Tran, E.J., 2016. Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation. Mol. Cell 61, 393–404. https://doi.org/10.1016/j.molcel.2015.12.024

Cohen, S., Puget, N., Lin, Y.-L., Clouaire, T., Aguirrebengoa, M., Rocher, V., Pasero, P., Canitrot, Y., Legube, G., 2018. Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. Nat. Commun. 9, 533. https://doi.org/10.1038/s41467-018-02894-w

Colak, D., Zaninovic, N., Cohen, M.S., Rosenwaks, Z., Yang, W.-Y., Gerhardt, J., Disney, M.D., Jaffrey, S.R., 2014. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. Science 343, 1002–1005. https://doi.org/10.1126/science.1245831

Costantino, L., Koshland, D., 2018. Genome-wide Map of R-Loop-Induced Damage Reveals How a Subset of R-Loops Contributes to Genomic Instability. Mol. Cell 71, 487-497.e3. https://doi.org/10.1016/j.molcel.2018.06.037

Costantino, L., Koshland, D., 2015. The Yin and Yang of R-loop biology. Curr. Opin. Cell Biol. 34, 39–45. https://doi.org/10.1016/j.ceb.2015.04.008

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., Melton, D.A., 2004. Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. 350, 1353–1356. https://doi.org/10.1056/NEJMsr040330

Cristini, A., Groh, M., Kristiansen, M.S., Gromak, N., 2018. RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. Cell Rep. 23, 1891–1905. https://doi.org/10.1016/j.celrep.2018.04.025

Cristini, A., Ricci, G., Britton, S., Salimbeni, S., Huang, S.-Y.N., Marinello, J., Calsou, P., Pommier, Y., Favre, G., Capranico, G., Gromak, N., Sordet, O., 2019. Dual Processing of R-Loops and Topoisomerase I Induces Transcription-Dependent DNA Double-Strand Breaks. Cell Rep. 28, 3167-3181.e6. https://doi.org/10.1016/j.celrep.2019.08.041

Crossley, M.P., Bocek, M., Cimprich, K.A., 2019. R-Loops as Cellular Regulators and Genomic Threats. Mol. Cell 73, 398–411. https://doi.org/10.1016/j.molcel.2019.01.024

Cummings, D.J., Tait, A., Goddard, J.M., 1974. Methylated bases in DNA from Paramecium aurelia. Biochim. Biophys. Acta 374, 1–11. https://doi.org/10.1016/0005-2787(74)90194-4

Dal Magro, C., Keller, P., Kotter, A., Werner, S., Duarte, V., Marchand, V., Ignarski, M., Freiwald, A., Müller, R.-U., Dieterich, C., Motorin, Y., Butter, F., Atta, M., Helm, M., 2018. A Vastly Increased Chemical Variety of RNA Modifications Containing a Thioacetal Structure. Angew. Chem. Int. Ed Engl. 57, 7893–7897. https://doi.org/10.1002/anie.201713188

Daniels, G.A., Lieber, M.R., 1995. RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. Nucleic Acids Res. 23, 5006–5011. https://doi.org/10.1093/nar/23.24.5006

De Magis, A., Manzo, S.G., Russo, M., Marinello, J., Morigi, R., Sordet, O., Capranico, G., 2019. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. Proc. Natl. Acad. Sci. U. S. A. 116, 816–825. https://doi.org/10.1073/pnas.1810409116

Deng, X., Chen, K., Luo, G.-Z., Weng, X., Ji, Q., Zhou, T., He, C., 2015. Widespread occurrence of N6-methyladenosine in bacterial mRNA. Nucleic Acids Res. 43, 6557–6567. https://doi.org/10.1093/nar/gkv596

Depamphilis, M.L., de Renty, C.M., Ullah, Z., Lee, C.Y., 2012. "The Octet": Eight Protein Kinases that Control Mammalian DNA Replication. Front. Physiol. 3, 368. https://doi.org/10.3389/fphys.2012.00368

Desrosiers, R., Friderici, K., Rottman, F., 1974. Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells. Proc. Natl. Acad. Sci. 71, 3971–3975. https://doi.org/10.1073/pnas.71.10.3971

Dodson, K.W., Berg, D.E., 1989. Factors affecting transposition activity of IS50 and Tn5 ends. Gene 76, 207–213. https://doi.org/10.1016/0378-1119(89)90161-3

Domínguez-Sánchez, M.S., Barroso, S., Gómez-González, B., Luna, R., Aguilera, A., 2011. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. PLoS Genet. 7, e1002386. https://doi.org/10.1371/journal.pgen.1002386

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., Rechavi, G., 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485, 201–206. https://doi.org/10.1038/nature11112

Douvlataniotis, K., Bensberg, M., Lentini, A., Gylemo, B., Nestor, C.E., 2020. No evidence for DNA N6-methyladenine in mammals. Sci. Adv. 6, eaay3335. https://doi.org/10.1126/sciadv.aay3335

Drolet, M., 2006. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. Mol. Microbiol. 59, 723–730. https://doi.org/10.1111/j.1365-2958.2005.05006.x

Drolet, M., Bi, X., Liu, L.F., 1994. Hypernegative supercoiling of the DNA template during transcription elongation in vitro. J. Biol. Chem. 269, 2068–2074.

Drolet, M., Phoenix, P., Menzel, R., Massé, E., Liu, L.F., Crouch, R.J., 1995. Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. Proc. Natl. Acad. Sci. U. S. A. 92, 3526–3530. https://doi.org/10.1073/pnas.92.8.3526

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., Wu, L., 2016. YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. 7, 12626. https://doi.org/10.1038/ncomms12626

Dumelie, J.G., Jaffrey, S.R., 2017. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. eLife 6. https://doi.org/10.7554/eLife.28306

Duncan, G., Firth, K., George, V., Hoang, M.D., Staniforth, A., Smith, G., Denning, C., 2017. Drug-Mediated Shortening of Action Potentials in LQTS2 Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes. Stem Cells Dev. 26, 1695–1705. https://doi.org/10.1089/scd.2017.0172

Dunn, D.B., Smith, J.D., 1955. Occurrence of a new base in the deoxyribonucleic acid of a strain of Bacterium coli. Nature 175, 336–337. https://doi.org/10.1038/175336a0

Dunn, K., Griffith, J.D., 1980. The presence of RNA in a double helix inhibits its interaction with histone protein. Nucleic Acids Res. 8, 555–566. https://doi.org/10.1093/nar/8.3.555

Dunnick, W., Hertz, G.Z., Scappino, L., Gritzmacher, C., 1993. DNA sequences at immunoglobulin switch region recombination sites. Nucleic Acids Res. 21, 365–372. https://doi.org/10.1093/nar/21.3.365

Ehrlich, M., Wilson, G.G., Kuo, K.C., Gehrke, C.W., 1987. N4-methylcytosine as a minor base in bacterial DNA. J. Bacteriol. 169, 939–943. https://doi.org/10.1128/jb.169.3.939-943.1987

Ekundayo, B., Bleichert, F., 2019. Origins of DNA replication. PLoS Genet. 15, e1008320. https://doi.org/10.1371/journal.pgen.1008320

El Hage, A., French, S.L., Beyer, A.L., Tollervey, D., 2010. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev. 24, 1546–1558. https://doi.org/10.1101/gad.573310

El Hage, A., Webb, S., Kerr, A., Tollervey, D., 2014. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet. 10, e1004716. https://doi.org/10.1371/journal.pgen.1004716

Engel, M., Eggert, C., Kaplick, P.M., Eder, M., Röh, S., Tietze, L., Namendorf, C., Arloth, J., Weber, P., Rex-Haffner, M., Geula, S., Jakovcevski, M., Hanna, J.H., Leshkowitz, D., Uhr, M., Wotjak, C.T., Schmidt, M.V., Deussing, J.M., Binder, E.B., Chen, A., 2018. The Role of m6A/m-RNA Methylation in Stress Response Regulation. Neuron 99, 389-403.e9. https://doi.org/10.1016/j.neuron.2018.07.009

Faget, L., Hnasko, T.S., 2015. Tyramide Signal Amplification for Immunofluorescent Enhancement, in: Hnasko, R. (Ed.), ELISA: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 161–172. https://doi.org/10.1007/978-1-4939-2742-5_16

Faili, A., Aoufouchi, S., Guéranger, Q., Zober, C., Léon, A., Bertocci, B., Weill, J.-C., Reynaud, C.-A., 2002. AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. Nat. Immunol. 3, 815–821. https://doi.org/10.1038/ni826

Fazi, F., Fatica, A., 2019. Interplay Between N6-Methyladenosine (m6A) and Non-coding RNAs in Cell Development and Cancer. Front. Cell Dev. Biol. 7, 116. https://doi.org/10.3389/fcell.2019.00116

Fei, Q., Zou, Z., Roundtree, I.A., Sun, H.-L., He, C., 2020. YTHDF2 promotes mitotic entry and is regulated by cell cycle mediators. PLoS Biol. 18, e3000664. https://doi.org/10.1371/journal.pbio.3000664

Finkel, D., Groner, Y., 1983. Methylations of adenosine residues (m6A) in pre-mRNA are important for formation of late simian virus 40 mRNAs. Virology 131, 409–425. https://doi.org/10.1016/0042-6822(83)90508-1

Flasch, D.A., Macia, Á., Sánchez, L., Ljungman, M., Heras, S.R., García-Pérez, J.L., Wilson, T.E., Moran, J.V., 2019. Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. Cell 177, 837-851.e28. https://doi.org/10.1016/j.cell.2019.02.050

Fragola, G., Mabb, A.M., Taylor-Blake, B., Niehaus, J.K., Chronister, W.D., Mao, H., Simon, J.M., Yuan, H., Li, Z., McConnell, M.J., Zylka, M.J., 2020. Deletion of Topoisomerase 1 in excitatory neurons causes genomic instability and early onset neurodegeneration. Nat. Commun. 11, 1962. https://doi.org/10.1038/s41467-020-15794-9

Frazee, A.C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., Leek, J.T., 2015. Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat. Biotechnol. 33, 243–246. https://doi.org/10.1038/nbt.3172

Fu, Y., Dominissini, D., Rechavi, G., He, C., 2014. Gene expression regulation mediated through reversible $m^6A$ RNA methylation. Nat. Rev. Genet. 15, 293–306. https://doi.org/10.1038/nrg3724

Fu, Y., Jia, G., Pang, X., Wang, R.N., Wang, X., Li, C.J., Smemo, S., Dai, Q., Bailey, K.A., Nobrega, M.A., Han, K.-L., Cui, Q., He, C., 2013. FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. Nat. Commun. 4, 1798. https://doi.org/10.1038/ncomms2822

Fu, Y., Luo, G.-Z., Chen, K., Deng, X., Yu, M., Han, D., Hao, Z., Liu, J., Lu, X., Dore, L.C., Weng, X., Ji, Q., Mets, L., He, C., 2015. N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas. Cell 161, 879–892. https://doi.org/10.1016/j.cell.2015.04.010

Fustin, J.-M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M.S., Kakeya, H., Manabe, I., Okamura, H., 2013. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell 155, 793–806. https://doi.org/10.1016/j.cell.2013.10.026

Gackowski, D., Starczak, M., Zarakowska, E., Modrzejewska, M., Szpila, A., Banaszkiewicz, Z., Olinski, R., 2016. Accurate, Direct, and High-Throughput Analyses of a Broad Spectrum of Endogenously Generated DNA Base Modifications with Isotope-Dilution Two-Dimensional Ultraperformance Liquid Chromatography with Tandem Mass Spectrometry: Possible Clinical Implication. Anal. Chem. 88, 12128–12136. https://doi.org/10.1021/acs.analchem.6b02900

Gaidamakov, S.A., Gorshkova, I.I., Schuck, P., Steinbach, P.J., Yamada, H., Crouch, R.J., Cerritelli, S.M., 2005. Eukaryotic RNases H1 act processively by interactions through the duplex RNA-binding domain. Nucleic Acids Res. 33, 2166–2175. https://doi.org/10.1093/nar/gki510

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., Li, X., 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes Dev. 25, 2041–2056. https://doi.org/10.1101/gad.17010011

García-Muse, T., Aguilera, A., 2019. R Loops: From Physiological to Pathological Roles. Cell 179, 604–618. https://doi.org/10.1016/j.cell.2019.08.055

García-Pichardo, D., Cañas, J.C., García-Rubio, M.L., Gómez-González, B., Rondón, A.G., Aguilera, A., 2017. Histone Mutants Separate R Loop Formation from Genome Instability Induction. Mol. Cell 66, 597-609.e5. https://doi.org/10.1016/j.molcel.2017.05.014

García-Rubio, M.L., Pérez-Calero, C., Barroso, S.I., Tumini, E., Herrera-Moyano, E., Rosado, I.V., Aguilera, A., 2015. The Fanconi Anemia Pathway Protects Genome Integrity from R-loops. PLOS Genet. 11, e1005674. https://doi.org/10.1371/journal.pgen.1005674

Georgoulis, A., Vorgias, C.E., Chrousos, G.P., Rogakou, E.P., 2017. Genome Instability and γH2AX. Int. J. Mol. Sci. 18. https://doi.org/10.3390/ijms18091979

Gerken, T., Girard, C.A., Tung, Y.-C.L., Webby, C.J., Saudek, V., Hewitson, K.S., Yeo, G.S.H., McDonough, M.A., Cunliffe, S., McNeill, L.A., Galvanovskis, J., Rorsman, P., Robins, P., Prieur, X., Coll, A.P., Ma, M., Jovanovic, Z., Farooqi, I.S., Sedgwick, B., Barroso, I., Lindahl, T., Ponting, C.P., Ashcroft, F.M., O'Rahilly, S., Schofield, C.J., 2007. The Obesity-Associated FTO Gene Encodes a 2-Oxoglutarate–Dependent Nucleic Acid Demethylase. Science 318, 1469–1472. https://doi.org/10.1126/science.1151710

Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., Ben-Haim, M.S., Eyal, E., Yunger, S., Pinto, Y., Jaitin, D.A., Viukov, S., Rais, Y., Krupalnik, V., Chomsky, E., Zerbib, M., Maza, I., Rechavi, Y., Massarwa, R., Hanna, S., Amit, I., Levanon, E.Y., Amariglio, N., Stern-Ginossar, N., Novershtern, N., Rechavi, G., Hanna, J.H., 2015. Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347, 1002–1006. https://doi.org/10.1126/science.1261417

Ginno, P.A., Burger, L., Seebacher, J., Iesmantavicius, V., Schübeler, D., 2018. Cell cycle-resolved chromatin proteomics reveals the extent of mitotic preservation of the genomic regulatory landscape. Nat. Commun. 9, 4048. https://doi.org/10.1038/s41467-018-06007-5

Ginno, P.A., Lim, Y.W., Lott, P.L., Korf, I., Chédin, F., 2013. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. 23, 1590–1600. https://doi.org/10.1101/gr.158436.113

Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I., Chédin, F., 2012. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell 45, 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Gogarten, J.P., Townsend, J.P., 2005. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. 3, 679–687. https://doi.org/10.1038/nrmicro1204

Gold, M., Hurwitz, J., Anders, M., 1963. The enzymatic methylation of RNA and DNA. Biochem. Biophys. Res. Commun. 11, 107–114. https://doi.org/10.1016/0006-291x(63)90075-5

Gómez-González, B., Felipe-Abrio, I., Aguilera, A., 2009. The S-Phase Checkpoint Is Required To Respond to R-Loops Accumulated in THO Mutants. Mol. Cell. Biol. 29, 5203–5213. https://doi.org/10.1128/MCB.00402-09

Gorovsky, M.A., Hattman, S., Pleger, G.L., 1973. ( 6 N)methyl adenine in the nuclear DNA of a eucaryote, Tetrahymena pyriformis. J. Cell Biol. 56, 697–701. https://doi.org/10.1083/jcb.56.3.697

Gorthi, A., Romero, J.C., Loranc, E., Cao, L., Lawrence, L.A., Goodale, E., Iniguez, A.B., Bernard, X., Masamsetti, V.P., Roston, S., Lawlor, E.R., Toretsky, J.A., Stegmaier, K., Lessnick, S.L., Chen, Y., Bishop, A.J.R., 2018. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. Nature 555, 387–391. https://doi.org/10.1038/nature25748

Graf, M., Bonetti, D., Lockhart, A., Serhal, K., Kellner, V., Maicher, A., Jolivet, P., Teixeira, M.T., Luke, B., 2017. Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. Cell 170, 72-85.e14. https://doi.org/10.1016/j.cell.2017.06.006

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., Shi, Y., 2015a. DNA Methylation on N6-Adenine in C. elegans. Cell 161, 868–878. https://doi.org/10.1016/j.cell.2015.04.005

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., Shi, Y., 2015b. DNA Methylation on N6-Adenine in C. elegans. Cell 161, 868–878. https://doi.org/10.1016/j.cell.2015.04.005

Greider, C.W., Blackburn, E.H., 1989. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature 337, 331–337. https://doi.org/10.1038/337331a0

Groh, M., Lufino, M.M.P., Wade-Martins, R., Gromak, N., 2014. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. PLoS Genet. 10, e1004318. https://doi.org/10.1371/journal.pgen.1004318

Gromak, N., West, S., Proudfoot, N.J., 2006. Pause sites promote transcriptional termination of mammalian RNA polymerase II. Mol. Cell. Biol. 26, 3986–3996. https://doi.org/10.1128/MCB.26.10.3986-3996.2006

Grosjean, H., 2009. Nucleic Acids Are Not Boring Long Polymers of Only Four Types of Nucleotides: A Guided Tour, in: DNA and RNA Modi Cation Enzymes: Structure, Mechanism, Function and Evolution. pp. 1–18.

Grunseich, C., Wang, I.X., Watts, J.A., Burdick, J.T., Guber, R.D., Zhu, Z., Bruzel, A., Lanman, T., Chen, K., Schindler, A.B., Edwards, N., Ray-Chaudhury, A., Yao, J., Lehky, T., Piszczek, G., Crain, B., Fischbeck, K.H., Cheung, V.G., 2018. Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters. Mol. Cell 69, 426-437.e7. https://doi.org/10.1016/j.molcel.2017.12.030

Gutiérrez, J.C., Callejas, S., Borniquel, S., Martín-González, A., 2000. DNA methylation in ciliates: implications in differentiation processes. Int. Microbiol. Off. J. Span. Soc. Microbiol. 3, 139–146.

Hainer, S.J., Pruneski, J.A., Mitchell, R.D., Monteverde, R.M., Martens, J.A., 2011. Intergenic transcription causes repression by directing nucleosome assembly. Genes Dev. 25, 29–40. https://doi.org/10.1101/gad.1975011

Halász, L., Karányi, Z., Boros-Oláh, B., Kuik-Rózsa, T., Sipos, É., Nagy, É., Mosolygó-L, Á., Mázló, A., Rajnavölgyi, É., Halmos, G., Székvölgyi, L., 2017. RNA-DNA hybrid (R-loop) immunoprecipitation mapping: an analytical workflow to evaluate inherent biases. Genome Res. 27, 1063–1073. https://doi.org/10.1101/gr.219394.116

Hamperl, S., Bocek, M.J., Saldivar, J.C., Swigut, T., Cimprich, K.A., 2017. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. Cell 170, 774-786.e19. https://doi.org/10.1016/j.cell.2017.07.043

Hao, Z., Wu, T., Cui, X., Zhu, P., Tan, C., Dou, X., Hsu, K.-W., Lin, Y.-T., Peng, P.-H., Zhang, L.-S., Gao, Y., Hu, L., Sun, H.-L., Zhu, A., Liu, J., Wu, K.-J., He, C., 2020. N6-Deoxyadenosine Methylation in Mammalian Mitochondrial DNA. Mol. Cell 78, 382-395.e8. https://doi.org/10.1016/j.molcel.2020.02.018

Harris, A.J., Goldman, A.D., 2020. The complex phylogenetic relationships of a 4mC/6mA DNA methyltransferase in prokaryotes. Mol. Phylogenet. Evol. 149, 106837. https://doi.org/10.1016/j.ympev.2020.106837

Hartono, S.R., Malapert, A., Legros, P., Bernard, P., Chédin, F., Vanoosthuyse, V., 2018. The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. J. Mol. Biol. 430, 272–284. https://doi.org/10.1016/j.jmb.2017.12.016

Hattman, S., Kenny, C., Berger, L., Pratt, K., 1978. Comparative study of DNA methylation in three unicellular eucaryotes. J. Bacteriol. 135, 1156–1157.

Haugland, R.A., Cline, M.G., 1980. Post-transcriptional modifications of oat coleoptile ribonucleic acids. 5'-Terminal capping and methylation of internal nucleosides in poly(A)-rich RNA. Eur. J. Biochem. 104, 271–277. https://doi.org/10.1111/j.1432-1033.1980.tb04425.x

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., Soller, M., 2016. m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature 540, 301–304. https://doi.org/10.1038/nature20577

He, C., 2010. Grand challenge commentary: RNA epigenetics? Nat. Chem. Biol. 6, 863–865. https://doi.org/10.1038/nchembio.482

Hegazy, Y.A., Fernando, C.M., Tran, E.J., 2020. The balancing act of R-loop biology: The good, the bad, and the ugly. J. Biol. Chem. 295, 905–913. https://doi.org/10.1074/jbc.REV119.011353

Hess, M.E., Hess, S., Meyer, K.D., Verhagen, L.A.W., Koch, L., Brönneke, H.S., Dietrich, M.O., Jordan, S.D., Saletore, Y., Elemento, O., Belgardt, B.F., Franz, T., Horvath, T.L., Rüther, U., Jaffrey, S.R., Kloppenburg, P., Brüning, J.C., 2013. The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat. Neurosci. 16, 1042–1048. https://doi.org/10.1038/nn.3449

Holliday, R., Pugh, J.E., 1975. DNA modification mechanisms and gene activity during development. Science 187, 226–232.
Hongay, C.F., Orr-Weaver, T.L., 2011. Drosophila Inducer of MEiosis 4 (IME4) is required for Notch signaling during oogenesis. Proc. Natl. Acad. Sci. U. S. A. 108, 14855–14860. https://doi.org/10.1073/pnas.1111577108
Horiuchi, K., Kawamura, T., Iwanari, H., Ohashi, R., Naito, M., Kodama, T., Hamakubo, T., 2013. Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. J. Biol. Chem. 288, 33292–33302. https://doi.org/10.1074/jbc.M113.500397
Hotchkiss, R.D., 1948. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J. Biol. Chem. 175, 315–332.
Hou, J., Zhang, He, Liu, J., Zhao, Zhenjun, Wang, J., Lu, Z., Hu, B., Zhou, J., Zhao, Zhicong, Feng, M., Zhang, Haiyan, Shen, B., Huang, X., Sun, B., He, C., Xia, Q., 2019. YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. Mol. Cancer 18, 163. https://doi.org/10.1186/s12943-019-1082-3
Hsieh, P., 2001. Molecular mechanisms of DNA mismatch repair. Mutat. Res. 486, 71–87. https://doi.org/10.1016/s0921-8777(01)00088-x
Hsu, P.J., Shi, H., He, C., 2017. Epitranscriptomic influences on development and disease. Genome Biol. 18, 197. https://doi.org/10.1186/s13059-017-1336-6
Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., Hu, Y.-C., Hüttelmaier, S., Skibbe, J.R., Su, R., Deng, X., Dong, L., Sun, M., Li, C., Nachtergaele, S., Wang, Y., Hu, C., Ferchen, K., Greis, K.D., Jiang, X., Wei, M., Qu, L., Guan, J.-L., He, C., Yang, J., Chen, J., 2018. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat. Cell Biol. 20, 285–295. https://doi.org/10.1038/s41556-018-0045-z
Huang, W., Xiong, J., Yang, Y., Liu, S.-M., Yuan, B.-F., Feng, Y.-Q., 2015. Determination of DNA adenine methylation in genomes of mammals and plants by liquid chromatography/mass spectrometry. RSC Adv. 5, 64046–64054. https://doi.org/10.1039/C5RA05307B
Huertas, P., Aguilera, A., 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell 12, 711–721. https://doi.org/10.1016/j.molcel.2003.08.010
Hulsen, T., de Vlieg, J., Alkema, W., 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9, 488. https://doi.org/10.1186/1471-2164-9-488
Hurwitz, J., Anders, M., Gold, M., Smith, I., 1965. THE ENZYMATIC METHYLATION OF RIBONUCLEIC ACID AND DEOXYRIBONUCLEIC ACID. VII. THE METHYLATION OF RIBONUCLEIC ACID. J. Biol. Chem. 240, 1256–1266.
Hyjek, M., Figiel, M., Nowotny, M., 2019. RNases H: Structure and mechanism. DNA Repair 84, 102672. https://doi.org/10.1016/j.dnarep.2019.102672
Imai, K., Slupphaug, G., Lee, W.-I., Revy, P., Nonoyama, S., Catalan, N., Yel, L., Forveille, M., Kavli, B., Krokan, H.E., Ochs, H.D., Fischer, A., Durandy, A., 2003. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat. Immunol. 4, 1023–1028. https://doi.org/10.1038/ni974
Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., Zhang, Y., 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129–1133. https://doi.org/10.1038/nature09303
Ito, S., Kuraoka, I., 2015. Epigenetic modifications in DNA could mimic oxidative DNA damage: A double-edged sword. DNA Repair 32, 52–57. https://doi.org/10.1016/j.dnarep.2015.04.013
Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 333, 1300–1303. https://doi.org/10.1126/science.1210597

Itoh, T., Tomizawa, J., 1980. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc. Natl. Acad. Sci. U. S. A. 77, 2450–2454. https://doi.org/10.1073/pnas.77.5.2450

Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P.N., Monahan, J., Carrieri, C., Enright, A.J., O'Carroll, D., 2017. The RNA m6A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence. Mol. Cell 67, 1059-1067.e4. https://doi.org/10.1016/j.molcel.2017.08.003

Iyer, L.M., Abhiman, S., Aravind, L., 2011. Natural history of eukaryotic DNA methylation systems. Prog. Mol. Biol. Transl. Sci. 101, 25–104. https://doi.org/10.1016/B978-0-12-387685-0.00002-0

Iyer, L.M., Zhang, D., Aravind, L., 2016. Adenine methylation in eukaryotes: Apprehending the complex evolutionary history and functional potential of an epigenetic modification. BioEssays News Rev. Mol. Cell. Dev. Biol. 38, 27–40. https://doi.org/10.1002/bies.201500104

Jenjaroenpun, P., Wongsurawat, T., Yenamandra, S.P., Kuznetsov, V.A., 2015. QmRLFS-finder: a model, web server and stand-alone tool for prediction and analysis of R-loop forming sequences. Nucleic Acids Res. 43, W527-534. https://doi.org/10.1093/nar/gkv344

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., He, C., 2011. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. 7, 885–887. https://doi.org/10.1038/nchembio.687

Jia, G., Yang, C.-G., Yang, S., Jian, X., Yi, C., Zhou, Z., He, C., 2008. Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. FEBS Lett. 582, 3313–3319. https://doi.org/10.1016/j.febslet.2008.08.019

Jiang, H., Lei, R., Ding, S.-W., Zhu, S., 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15, 182. https://doi.org/10.1186/1471-2105-15-182

Jimeno, S., Rondón, A.G., Luna, R., Aguilera, A., 2002. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. EMBO J. 21, 3526–3535. https://doi.org/10.1093/emboj/cdf335

Kabeche, L., Nguyen, H.D., Buisson, R., Zou, L., 2018. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 359, 108–114. https://doi.org/10.1126/science.aan6490

Kafer, G.R., Li, X., Horii, T., Suetake, I., Tajima, S., Hatada, I., Carlton, P.M., 2016. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. Cell Rep. 14, 1283–1292. https://doi.org/10.1016/j.celrep.2016.01.035

Kan, L., Grozhik, A.V., Vedanayagam, J., Patil, D.P., Pang, N., Lim, K.-S., Huang, Y.-C., Joseph, B., Lin, C.-J., Despic, V., Guo, J., Yan, D., Kondo, S., Deng, W.-M., Dedon, P.C., Jaffrey, S.R., Lai, E.C., 2017. The m6A pathway facilitates sex determination in Drosophila. Nat. Commun. 8, 15737. https://doi.org/10.1038/ncomms15737

Kao, Y.-P., Hsieh, W.-C., Hung, S.-T., Huang, C.-W., Lieber, M.R., Huang, F.-T., 2013. Detection and characterization of R-loops at the murine immunoglobulin Sα region. Mol. Immunol. 54, 208–216. https://doi.org/10.1016/j.molimm.2012.11.009

Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., Wang, P.J., 2018. Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. PLoS Genet. 14, e1007412. https://doi.org/10.1371/journal.pgen.1007412

Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., Vågbø, C.B., Kusśnierczyk, A., Klungland, A., Darnell, J.E., Darnell, R.B., 2015. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. Genes Dev. 29, 2037–2053. https://doi.org/10.1101/gad.269415.115

Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vågbø, C.B., Geula, S., Hanna, J.H., Black, D.L., Darnell, J.E., Darnell, R.B., 2017. m6A mRNA modifications are deposited in nascent pre-mRNA and are

not required for splicing but do specify cytoplasmic turnover. Genes Dev. 31, 990–1006. https://doi.org/10.1101/gad.301036.117
Kim, H.D., Choe, J., Seo, Y.S., 1999. The sen1(+) gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. Biochemistry 38, 14697–14710. https://doi.org/10.1021/bi991470c
Kim, N., Jinks-Robertson, S., 2012. Transcription as a source of genome instability. Nat. Rev. Genet. 13, 204–214. https://doi.org/10.1038/nrg3152
Klawitter, S., Fuchs, N.V., Upton, K.R., Muñoz-Lopez, M., Shukla, R., Wang, J., Garcia-Cañadas, M., Lopez-Ruiz, C., Gerhardt, D.J., Sebe, A., Grabundzija, I., Merkert, S., Gerdes, P., Pulgarin, J.A., Bock, A., Held, U., Witthuhn, A., Haase, A., Sarkadi, B., Löwer, J., Wolvetang, E.J., Martin, U., Ivics, Z., Izsvák, Z., Garcia-Perez, J.L., Faulkner, G.J., Schumann, G.G., 2016. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat. Commun. 7, 10286. https://doi.org/10.1038/ncomms10286
Koh, C.W.Q., Goh, Y.T., Goh, W.S.S., 2019. Atlas of quantitative single-base-resolution N6-methyl-adenine methylomes. Nat. Commun. 10, 5636. https://doi.org/10.1038/s41467-019-13561-z
Kotsantis, P., Silva, L.M., Irmscher, S., Jones, R.M., Folkes, L., Gromak, N., Petermann, E., 2016. Increased global transcription activity as a mechanism of replication stress in cancer. Nat. Commun. 7, 13087. https://doi.org/10.1038/ncomms13087
Koziol, M.J., Bradshaw, C.R., Allen, G.E., Costa, A.S.H., Frezza, C., Gurdon, J.B., 2016. Identification of methylated deoxyadenosines in vertebrates reveals diversity in DNA modifications. Nat. Struct. Mol. Biol. 23, 24–30. https://doi.org/10.1038/nsmb.3145
Kretschmer, J., Rao, H., Hackert, P., Sloan, K.E., Höbartner, C., Bohnsack, M.T., 2018. The m6A reader protein YTHDC2 interacts with the small ribosomal subunit and the 5'-3' exoribonuclease XRN1. RNA N. Y. N 24, 1339–1350. https://doi.org/10.1261/rna.064238.117
Krug, R.M., Morgan, M.A., Shatkin, A.J., 1976. Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. J. Virol. 20, 45–53.
Kurowski, M.A., Bhagwat, A.S., Papaj, G., Bujnicki, J.M., 2003. Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. BMC Genomics 4, 48. https://doi.org/10.1186/1471-2164-4-48
Kweon, S.-M., Chen, Y., Moon, E., Kvederaviciutė, K., Klimasauskas, S., Feldman, D.E., 2019. An Adversarial DNA N6-Methyladenine-Sensor Network Preserves Polycomb Silencing. Mol. Cell 74, 1138-1147.e6. https://doi.org/10.1016/j.molcel.2019.03.018
Lang, K.S., Hall, A.N., Merrikh, C.N., Ragheb, M., Tabakh, H., Pollock, A.J., Woodward, J.J., Dreifus, J.E., Merrikh, H., 2017. Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. Cell 170, 787-799.e18. https://doi.org/10.1016/j.cell.2017.07.044
Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. https://doi.org/10.1186/gb-2009-10-3-r25
Lasman, L., Krupalnik, V., Geula, S., Zerbib, M., Viukov, S., Mor, N., Castrejon, A.A., Mizrahi, O., Shashank, S., Nachshon, A., Schneir, D., Aigner, S., Shankar, A., Mueller, J., Stern-Ginossar, N., Yeo, G.W., Novershtern, N., Hanna, J.H., 2020. Context-dependent functional compensation between Ythdf m6A readers. bioRxiv 2020.06.03.131441. https://doi.org/10.1101/2020.06.03.131441
Lee, H., Bao, S., Qian, Y., Geula, S., Leslie, J., Zhang, C., Hanna, J.H., Ding, L., 2019. Stage-specific requirement for Mettl3-dependent m6A mRNA methylation during haematopoietic stem cell differentiation. Nat. Cell Biol. 21, 700–709. https://doi.org/10.1038/s41556-019-0318-1
Lee, Y., Choe, J., Park, O.H., Kim, Y.K., 2020. Molecular Mechanisms Driving mRNA Degradation by m6A Modification. Trends Genet. TIG 36, 177–188. https://doi.org/10.1016/j.tig.2019.12.007

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., Roignant, J.-Y., 2016. m6A modulates neuronal functions and sex determination in Drosophila. Nature 540, 242–247. https://doi.org/10.1038/nature20568

Lentini, A., Lagerwall, C., Vikingsson, S., Mjoseng, H.K., Douvlataniotis, K., Vogt, H., Green, H., Meehan, R.R., Benson, M., Nestor, C.E., 2018. A reassessment of DNA-immunoprecipitation-based genomic profiling. Nat. Methods 15, 499–504. https://doi.org/10.1038/s41592-018-0038-7

Lesbirel, S., Viphakone, N., Parker, M., Parker, J., Heath, C., Sudbery, I., Wilson, S.A., 2018. The m6A-methylase complex recruits TREX and regulates mRNA export. Sci. Rep. 8, 13827. https://doi.org/10.1038/s41598-018-32310-8

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma. Oxf. Engl. 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., Perez, S.P., Suganthan, R., He, C., Bjørås, M., Klungland, A., 2018. Ythdf2-mediated m6A mRNA clearance modulates neural development in mice. Genome Biol. 19, 69. https://doi.org/10.1186/s13059-018-1436-y

Li, X., Manley, J.L., 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122, 365–378. https://doi.org/10.1016/j.cell.2005.06.008

Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., Huang, Huilin, Nachtergaele, S., Dong, L., Hu, C., Qin, X., Tang, L., Wang, Y., Hong, G.-M., Huang, Hao, Wang, X., Chen, P., Gurbuxani, S., Arnovitz, S., Li, Y., Li, S., Strong, J., Neilly, M.B., Larson, R.A., Jiang, X., Zhang, P., Jin, J., He, C., Chen, J., 2017. FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N6-Methyladenosine RNA Demethylase. Cancer Cell 31, 127–141. https://doi.org/10.1016/j.ccell.2016.11.017

Liao, S., Sun, H., Xu, C., 2018. YTH Domain: A Family of N6-methyladenosine (m6A) Readers. Genomics Proteomics Bioinformatics 16, 99–107. https://doi.org/10.1016/j.gpb.2018.04.002

Lim, Y.W., Sanz, L.A., Xu, X., Hartono, S.R., Chédin, F., 2015. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutières syndrome. eLife 4. https://doi.org/10.7554/eLife.08007

Lin, S., Choe, J., Du, P., Triboulet, R., Gregory, R.I., 2016. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol. Cell 62, 335–345. https://doi.org/10.1016/j.molcel.2016.03.021

Lin, X., Chai, G., Wu, Y., Li, J., Chen, F., Liu, J., Luo, G., Tauler, J., Du, J., Lin, S., He, C., Wang, H., 2019. RNA m6A methylation regulates the epithelial mesenchymal transition of cancer cells and translation of Snail. Nat. Commun. 10, 2065. https://doi.org/10.1038/s41467-019-09865-9

Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362, 709–715. https://doi.org/10.1038/362709a0

Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., Jaffrey, S.R., 2015. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat. Methods 12, 767–772. https://doi.org/10.1038/nmeth.3453

Little, N.A., Hastie, N.D., Davies, R.C., 2000. Identification of WTAP, a novel Wilms' tumour 1-associating protein. Hum. Mol. Genet. 9, 2231–2239. https://doi.org/10.1093/oxfordjournals.hmg.a018914

Liu, J., Dou, X., Chen, Chuanyuan, Chen, Chuan, Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., He, C., 2020. N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science 367, 580–586. https://doi.org/10.1126/science.aay6018

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W., He, C., 2014. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. 10, 93–95. https://doi.org/10.1038/nchembio.1432

Liu, J., Zhu, Y., Luo, G.-Z., Wang, Xinxia, Yue, Y., Wang, Xiaona, Zong, X., Chen, K., Yin, H., Fu, Y., Han, D., Wang, Y., Chen, D., He, C., 2016. Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. Nat. Commun. 7, 13052. https://doi.org/10.1038/ncomms13052

Liu, L.F., Wang, J.C., 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. U. S. A. 84, 7024–7027. https://doi.org/10.1073/pnas.84.20.7024

Liu, N., Zhou, K.I., Parisien, M., Dai, Q., Diatchenko, L., Pan, T., 2017. N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res. 45, 6051–6063. https://doi.org/10.1093/nar/gkx141

Liu, T., Wei, Q., Jin, J., Luo, Q., Liu, Y., Yang, Y., Cheng, C., Li, Lanfang, Pi, J., Si, Y., Xiao, H., Li, Li, Rao, S., Wang, F., Yu, Jianhua, Yu, Jia, Zou, D., Yi, P., 2020. The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. Nucleic Acids Res. 48, 3816–3831. https://doi.org/10.1093/nar/gkaa048

Liu, X., Lai, W., Li, Y., Chen, S., Liu, B., Zhang, N., Mo, J., Lyu, C., Zheng, J., Du, Y.-R., Jiang, G., Xu, G.-L., Wang, H., 2020. N6-methyladenine is incorporated into mammalian genome by DNA polymerase. Cell Res. https://doi.org/10.1038/s41422-020-0317-6

Liu, X.-M., Zhou, J., Mao, Y., Ji, Q., Qian, S.-B., 2019. Programmable RNA N6-methyladenosine editing by CRISPR-Cas9 conjugates. Nat. Chem. Biol. 15, 865–871. https://doi.org/10.1038/s41589-019-0327-1

Lluch-Senar, M., Luong, K., Lloréns-Rico, V., Delgado, J., Fang, G., Spittle, K., Clark, T.A., Schadt, E., Turner, S.W., Korlach, J., Serrano, L., 2013. Comprehensive methylome characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. PLoS Genet. 9, e1003191. https://doi.org/10.1371/journal.pgen.1003191

Lockhart, A., Pires, V.B., Bento, F., Kellner, V., Luke-Glaser, S., Yakoub, G., Ulrich, H.D., Luke, B., 2019. RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids. Cell Rep. 29, 2890-2900.e5. https://doi.org/10.1016/j.celrep.2019.10.108

Lombraña, R., Almeida, R., Álvarez, A., Gómez, M., 2015. R-loops and initiation of DNA replication in human cells: a missing link? Front. Genet. 6, 158. https://doi.org/10.3389/fgene.2015.00158

Low, D.A., Weyand, N.J., Mahan, M.J., 2001. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. Infect. Immun. 69, 7197–7204. https://doi.org/10.1128/IAI.69.12.7197-7204.2001

Lu, M., Campbell, J.L., Boye, E., Kleckner, N., 1994. SeqA: a negative modulator of replication initiation in E. coli. Cell 77, 413–426. https://doi.org/10.1016/0092-8674(94)90156-2

Lu, W.-T., Hawley, B.R., Skalka, G.L., Baldock, R.A., Smith, E.M., Bader, A.S., Malewicz, M., Watts, F.Z., Wilczynska, A., Bushell, M., 2018. Drosha drives the formation of DNA:RNA hybrids around DNA break sites to facilitate DNA repair. Nat. Commun. 9, 532. https://doi.org/10.1038/s41467-018-02893-x

Lujan, S.A., Williams, J.S., Clausen, A.R., Clark, A.B., Kunkel, T.A., 2013. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. Mol. Cell 50, 437–443. https://doi.org/10.1016/j.molcel.2013.03.017

Luna, R., Jimeno, S., Marín, M., Huertas, P., García-Rubio, M., Aguilera, A., 2005. Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. Mol. Cell 18, 711–722. https://doi.org/10.1016/j.molcel.2005.05.001

Lyon, M.F., 1961. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372–373. https://doi.org/10.1038/190372a0

Machnicka, M.A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S., Dunin-Horkawicz, S., Rother, K.M., Helm, M., Bujnicki, J.M.,

Grosjean, H., 2013. MODOMICS: a database of RNA modification pathways--2013 update. Nucleic Acids Res. 41, D262-267. https://doi.org/10.1093/nar/gks1007
Maden, B.E., 1986. Identification of the locations of the methyl groups in 18 S ribosomal RNA from Xenopus laevis and man. J. Mol. Biol. 189, 681–699. https://doi.org/10.1016/0022-2836(86)90498-5
Malig, M., Hartono, S.R., Giafaglione, J.M., Sanz, L.A., Chedin, F., 2020. Ultra-deep Coverage Single-molecule R-loop Footprinting Reveals Principles of R-loop Formation. J. Mol. Biol. 432, 2271–2288. https://doi.org/10.1016/j.jmb.2020.02.014
Manis, J.P., Tian, M., Alt, F.W., 2002. Mechanism and control of class-switch recombination. Trends Immunol. 23, 31–39. https://doi.org/10.1016/s1471-4906(01)02111-1
Marinov, G.K., Kundaje, A., Park, P.J., Wold, B.J., 2014. Large-scale quality analysis of published ChIP-seq data. G3 Bethesda Md 4, 209–223. https://doi.org/10.1534/g3.113.008680
Masukata, H., Dasgupta, S., Tomizawa, J., 1987. Transcriptional activation of ColE1 DNA synthesis by displacement of the nontranscribed strand. Cell 51, 1123–1130. https://doi.org/10.1016/0092-8674(87)90598-8
Matos, D.A., Zhang, J.-M., Ouyang, J., Nguyen, H.D., Genois, M.-M., Zou, L., 2020. ATR Protects the Genome against R Loops through a MUS81-Triggered Feedback Loop. Mol. Cell 77, 514-527.e4. https://doi.org/10.1016/j.molcel.2019.10.010
Mattick, J.S., 2018. The State of Long Non-Coding RNA Biology. Non-Coding RNA 4. https://doi.org/10.3390/ncrna4030017
Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.-J., Chen, Q., Gross, S.S., Elemento, O., Debart, F., Kiledjian, M., Jaffrey, S.R., 2017. Reversible methylation of m6Am in the 5’ cap controls mRNA stability. Nature 541, 371–375. https://doi.org/10.1038/nature21022
Mendel, M., Chen, K.-M., Homolka, D., Gos, P., Pandey, R.R., McCarthy, A.A., Pillai, R.S., 2018. Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development. Mol. Cell 71, 986-1000.e11. https://doi.org/10.1016/j.molcel.2018.08.004
Meng, T.-G., Lu, X., Guo, L., Hou, G.-M., Ma, X.-S., Li, Q.-N., Huang, L., Fan, L.-H., Zhao, Z.-H., Ou, X.-H., OuYang, Y.-C., Schatten, H., Li, L., Wang, Z.-B., Sun, Q.-Y., 2019. Mettl14 is required for mouse postimplantation development by facilitating epiblast maturation. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 33, 1179–1187. https://doi.org/10.1096/fj.201800719R
Meyer, K.D., Jaffrey, S.R., 2014. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat. Rev. Mol. Cell Biol. 15, 313–326. https://doi.org/10.1038/nrm3785
Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., Jaffrey, S.R., 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3’ UTRs and near stop codons. Cell 149, 1635–1646. https://doi.org/10.1016/j.cell.2012.05.003
Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., Proudfoot, N.J., 2011. Yeast Sen1 Helicase Protects the Genome from Transcription-Associated Instability. Mol. Cell 41, 21–32. https://doi.org/10.1016/j.molcel.2010.12.007
Mita, P., Wudzinska, A., Sun, X., Andrade, J., Nayak, S., Kahler, D.J., Badri, S., LaCava, J., Ueberheide, B., Yun, C.Y., Fenyö, D., Boeke, J.D., n.d. LINE-1 protein localization and functional dynamics during the cell cycle. eLife 7. https://doi.org/10.7554/eLife.30058
Mohandas, T., Sparkes, R.S., Shapiro, L.J., 1981. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science 211, 393–396. https://doi.org/10.1126/science.6164095
Mondo, S.J., Dannebaum, R.O., Kuo, R.C., Louie, K.B., Bewick, A.J., LaButti, K., Haridas, S., Kuo, A., Salamov, A., Ahrendt, S.R., Lau, R., Bowen, B.P., Lipzen, A., Sullivan, W., Andreopoulos, B.B., Clum, A., Lindquist, E., Daum, C., Northen, T.R., Kunde-Ramamoorthy, G., Schmitz, R.J., Gryganskyi, A., Culley, D., Magnuson, J., James, T.Y., O’Malley, M.A., Stajich, J.E., Spatafora,

J.W., Visel, A., Grigoriev, I.V., 2017. Widespread adenine N6-methylation of active genes in fungi. Nat. Genet. 49, 964–968. https://doi.org/10.1038/ng.3859
Müller, S., Glaß, M., Singh, A.K., Haase, J., Bley, N., Fuchs, T., Lederer, M., Dahl, A., Huang, H., Chen, J., Posern, G., Hüttelmaier, S., 2019. IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. Nucleic Acids Res. 47, 375–390. https://doi.org/10.1093/nar/gky1012
Nadel, J., Athanasiadou, R., Lemetre, C., Wijetunga, N.A., Ó Broin, P., Sato, H., Zhang, Z., Jeddeloh, J., Montagna, C., Golden, A., Seoighe, C., Greally, J.M., 2015. RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin 8, 46. https://doi.org/10.1186/s13072-015-0040-6
Nair, L., Chung, H., Basu, U., 2020. Regulation of long non-coding RNAs and genome dynamics by the RNA surveillance machinery. Nat. Rev. Mol. Cell Biol. 21, 123–136. https://doi.org/10.1038/s41580-019-0209-0
Nance, D.J., Satterwhite, E.R., Bhaskar, B., Misra, S., Carraway, K.R., Mansfield, K.D., 2020. Characterization of METTL16 as a cytoplasmic RNA binding protein. PloS One 15, e0227647. https://doi.org/10.1371/journal.pone.0227647
Narayan, P., Rottman, F.M., 1988. An in vitro system for accurate methylation of internal adenosine residues in messenger RNA. Science 242, 1159–1162. https://doi.org/10.1126/science.3187541
Nguyen, H.D., Leong, W.Y., Li, W., Reddy, P.N.G., Sullivan, J.D., Walter, M.J., Zou, L., Graubert, T.A., 2018. Spliceosome Mutations Induce R Loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndromes. Cancer Res. 78, 5363–5374. https://doi.org/10.1158/0008-5472.CAN-17-3970
Nguyen, H.D., Zou, L., Graubert, T.A., 2019. Targeting R-loop-associated ATR response in myelodysplastic syndrome. Oncotarget 10, 2581–2582. https://doi.org/10.18632/oncotarget.26851
Nick McElhinny, S.A., Watts, B.E., Kumar, D., Watt, D.L., Lundström, E.-B., Burgers, P.M.J., Johansson, E., Chabes, A., Kunkel, T.A., 2010. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. Proc. Natl. Acad. Sci. U. S. A. 107, 4949–4954. https://doi.org/10.1073/pnas.0914857107
Niehrs, C., Luke, B., 2020. Regulatory R-loops as facilitators of gene expression and genome stability. Nat. Rev. Mol. Cell Biol. 21, 167–178. https://doi.org/10.1038/s41580-019-0206-3
Nikolskaya, I.I., Lopatina, N.G., Debov, S.S., 1981. On heterogeneity of DNA methylases from Escherichia coli SK cells. Mol. Cell. Biochem. 35, 3–10. https://doi.org/10.1007/BF02358182
Niu, Y., Lin, Z., Wan, A., Chen, H., Liang, H., Sun, L., Wang, Y., Li, X., Xiong, X.-F., Wei, B., Wu, X., Wan, G., 2019. RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. Mol. Cancer 18, 46. https://doi.org/10.1186/s12943-019-1004-4
Nudler, E., Mustaev, A., Lukhtanov, E., Goldfarb, A., 1997. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell 89, 33–41. https://doi.org/10.1016/s0092-8674(00)80180-4
O'Brown, Z.K., Boulias, K., Wang, J., Wang, S.Y., O'Brown, N.M., Hao, Z., Shibuya, H., Fady, P.-E., Shi, Y., He, C., Megason, S.G., Liu, T., Greer, E.L., 2019. Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. BMC Genomics 20, 445. https://doi.org/10.1186/s12864-019-5754-6
O'Brown, Z.K., Greer, E.L., 2016. N6-Methyladenine: A Conserved and Dynamic DNA Mark. Adv. Exp. Med. Biol. 945, 213–246. https://doi.org/10.1007/978-3-319-43624-1_10
Ohle, C., Tesorero, R., Schermann, G., Dobrev, N., Sinning, I., Fischer, T., 2016. Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. Cell 167, 1001-1013.e7. https://doi.org/10.1016/j.cell.2016.10.001

Okano, M., Bell, D.W., Haber, D.A., Li, E., 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99, 247–257. https://doi.org/10.1016/s0092-8674(00)81656-6

Olinski, R., Starczak, M., Gackowski, D., 2016. Enigmatic 5-hydroxymethyluracil: Oxidatively modified base, epigenetic mark or both? Mutat. Res. Rev. Mutat. Res. 767, 59–66. https://doi.org/10.1016/j.mrrev.2016.02.001

Palozola, K.C., Donahue, G., Liu, H., Grant, G.R., Becker, J.S., Cote, A., Yu, H., Raj, A., Zaret, K.S., 2017. Mitotic Transcription and Waves of Gene Reactivation During Mitotic Exit. Science 358, 119–122. https://doi.org/10.1126/science.aal4671

Paris, J., Morgan, M., Campos, J., Spencer, G.J., Shmakova, A., Ivanova, I., Mapperley, C., Lawson, H., Wotherspoon, D.A., Sepulveda, C., Vukovic, M., Allen, L., Sarapuu, A., Tavosanis, A., Guitart, A.V., Villacreces, A., Much, C., Choe, J., Azar, A., van de Lagemaat, L.N., Vernimmen, D., Nehme, A., Mazurier, F., Somervaille, T.C.P., Gregory, R.I., O'Carroll, D., Kranc, K.R., 2019. Targeting the RNA m6A Reader YTHDF2 Selectively Compromises Cancer Stem Cells in Acute Myeloid Leukemia. Cell Stem Cell 25, 137-148.e6. https://doi.org/10.1016/j.stem.2019.03.021

Park, O.H., Ha, H., Lee, Y., Boo, S.H., Kwon, D.H., Song, H.K., Kim, Y.K., 2019. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. Mol. Cell 74, 494-507.e8. https://doi.org/10.1016/j.molcel.2019.02.034

Patil, D.P., Chen, C.-K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M., Jaffrey, S.R., 2016. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature 537, 369–373. https://doi.org/10.1038/nature19342

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., Meyer, T., Cimprich, K.A., 2009. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell 35, 228–239. https://doi.org/10.1016/j.molcel.2009.06.021

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N., Basu, U., 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature 514, 389–393. https://doi.org/10.1038/nature13580

Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., Conrad, N.K., 2017. The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell 169, 824-835.e14. https://doi.org/10.1016/j.cell.2017.05.003

Percharde, M., Lin, C.-J., Yin, Y., Guan, J., Peixoto, G.A., Bulut-Karslioglu, A., Biechele, S., Huang, B., Shen, X., Ramalho-Santos, M., 2018. A LINE1-Nucleolin partnership regulates early development and ES cell identity. Cell 174, 391-405.e19. https://doi.org/10.1016/j.cell.2018.05.043

Perry, R.P., Kelley, D.E., 1975. Methylated constituents of heterogeneous nuclear RNA: presence in blocked 5' terminal structures. Cell 6, 13–19. https://doi.org/10.1016/0092-8674(75)90068-9

Petersen-Mahrt, S.K., Harris, R.S., Neuberger, M.S., 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature 418, 99–103. https://doi.org/10.1038/nature00862

Pfaffeneder, T., Hackner, B., Truss, M., Münzel, M., Müller, M., Deiml, C.A., Hagemeier, C., Carell, T., 2011. The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew. Chem. Int. Ed Engl. 50, 7008–7012. https://doi.org/10.1002/anie.201103899

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45. https://doi.org/10.1093/nar/29.9.e45

Pheasant, M., Mattick, J.S., 2007. Raising the estimate of functional human sequences. Genome Res. 17, 1245–1253. https://doi.org/10.1101/gr.6406307

Phillips, D.D., Garboczi, D.N., Singh, K., Hu, Z., Leppla, S.H., Leysath, C.E., 2013. The sub-nanomolar binding of DNA-RNA hybrids by the single-chain Fv fragment of antibody S9.6. J. Mol. Recognit. JMR 26, 376–381. https://doi.org/10.1002/jmr.2284

Ping, X.-L., Sun, B.-F., Wang, L., Xiao, W., Yang, X., Wang, W.-J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.-S., Zhao, X., Li, A., Yang, Y., Dahal, U., Lou, X.-M., Liu, X., Huang, J., Yuan, W.-P., Zhu, X.-F., Cheng, T., Zhao, Y.-L., Wang, X., Rendtlew Danielsen, J.M., Liu, F., Yang, Y.-G., 2014. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 24, 177–189. https://doi.org/10.1038/cr.2014.3

Pohjoismäki, J.L.O., Holmes, J.B., Wood, S.R., Yang, M.-Y., Yasukawa, T., Reyes, A., Bailey, L.J., Cluett, T.J., Goffart, S., Willcox, S., Rigby, R.E., Jackson, A.P., Spelbrink, J.N., Griffith, J.D., Crouch, R.J., Jacobs, H.T., Holt, I.J., 2010. Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of RNA/DNA hybrid. J. Mol. Biol. 397, 1144–1155. https://doi.org/10.1016/j.jmb.2010.02.029

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma. Oxf. Engl. 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., Amit, I., Regev, A., 2011. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. 29, 436–442. https://doi.org/10.1038/nbt.1861

Rae, P.M.M., Spear, B.B., 1978. Macronuclear DNA of the hypotrichous ciliate Oxytricha fallax. Proc. Natl. Acad. Sci. U. S. A. 75, 4992–4996.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160-165. https://doi.org/10.1093/nar/gkw257

Ramsawhook, A.H., Lewis, L.C., Eleftheriou, M., Abakir, A., Durczak, P., Markus, R., Rajani, S., Hannan, N.R.F., Coyle, B., Ruzov, A., 2017. Immunostaining for DNA Modifications: Computational Analysis of Confocal Images. J. Vis. Exp. JoVE. https://doi.org/10.3791/56318

Ratel, D., Ravanat, J.-L., Berger, F., Wion, D., 2006. N6-methyladenine: the other methylated base of DNA. BioEssays News Rev. Mol. Cell. Dev. Biol. 28, 309–315. https://doi.org/10.1002/bies.20342

Reaban, M.E., Griffin, J.A., 1990. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. Nature 348, 342–344. https://doi.org/10.1038/348342a0

Reijns, M.A.M., Rabe, B., Rigby, R.E., Mill, P., Astell, K.R., Lettice, L.A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P.S., Sexton, D., Grimes, G., Holt, I.J., Hill, R.E., Taylor, M.S., Lawson, K.A., Dorin, J.R., Jackson, A.P., 2012. Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. Cell 149, 1008–1022. https://doi.org/10.1016/j.cell.2012.04.011

Reik, W., Collick, A., Norris, M.L., Barton, S.C., Surani, M.A., 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. Nature 328, 248–251. https://doi.org/10.1038/328248a0

Reisenauer, A., Kahng, L.S., McCollum, S., Shapiro, L., 1999. Bacterial DNA methylation: a cell cycle regulator? J. Bacteriol. 181, 5135–5139.

Richard, P., Manley, J.L., 2017. R Loops and Links to Human Disease. J. Mol. Biol. 429, 3168–3180. https://doi.org/10.1016/j.jmb.2016.08.031

Rieder, C.L., Palazzo, R.E., 1992. Colcemid and the mitotic cycle. J. Cell Sci. 102 ( Pt 3), 387–392.

Roberts, D., Hoopes, B.C., McClure, W.R., Kleckner, N., 1985. IS10 transposition is regulated by DNA adenine methylation. Cell 43, 117–130. https://doi.org/10.1016/0092-8674(85)90017-0

Roberts, R.W., Crothers, D.M., 1992. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. Science 258, 1463–1466. https://doi.org/10.1126/science.1279808

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., Ali, S., Chin, S.-F., Palmieri, C., Caldas, C., Carroll, J.S., 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481, 389–393. https://doi.org/10.1038/nature10730

Rottman, F.M., Bokar, J.A., Narayan, P., Shambaugh, M.E., Ludwiczak, R., 1994. N6-adenosine methylation in mRNA: substrate specificity and enzyme complexity. Biochimie 76, 1109–1114. https://doi.org/10.1016/0300-9084(94)90038-8

Roundtree, I.A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., He, E., Shen, B., He, C., 2017. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. eLife 6. https://doi.org/10.7554/eLife.31311

Roy, D., Zhang, Z., Lu, Z., Hsieh, C.-L., Lieber, M.R., 2010. Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site. Mol. Cell. Biol. 30, 146–159. https://doi.org/10.1128/MCB.00897-09

Salnikow, J., Moore, S., Stein, W.H., 1970. Comparison of the multiple forms of bovine pancreatic deoxyribonuclease. J. Biol. Chem. 245, 5685–5690.

Santos, F., Dean, W., 2006. Using immunofluorescence to observe methylation changes in mammalian preimplantation embryos. Methods Mol. Biol. Clifton NJ 325, 129–137. https://doi.org/10.1385/1-59745-005-7:129

Santos-Pereira, J.M., Aguilera, A., 2015. R loops: new modulators of genome dynamics and function. Nat. Rev. Genet. 16, 583–597. https://doi.org/10.1038/nrg3961

Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X., Chédin, F., 2016a. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol. Cell 63, 167–178. https://doi.org/10.1016/j.molcel.2016.05.032

Sanz, L.A., Hartono, S.R., Lim, Y.W., Steyaert, S., Rajpurkar, A., Ginno, P.A., Xu, X., Chédin, F., 2016b. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol. Cell 63, 167–178. https://doi.org/10.1016/j.molcel.2016.05.032

Saponaro, M., Kantidakis, T., Mitter, R., Kelly, G.P., Heron, M., Williams, H., Söding, J., Stewart, A., Svejstrup, J.Q., 2014. RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. Cell 157, 1037–1049. https://doi.org/10.1016/j.cell.2014.03.048

Schibler, U., Kelley, D.E., Perry, R.P., 1977. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J. Mol. Biol. 115, 695–714. https://doi.org/10.1016/0022-2836(77)90110-3

Schiffers, S., Ebert, C., Rahimoff, R., Kosmatchev, O., Steinbacher, J., Bohne, A.-V., Spada, F., Michalakis, S., Nickelsen, J., Müller, M., Carell, T., 2017. Quantitative LC-MS Provides No Evidence for m6 dA or m4 dC in the Genome of Mouse Embryonic Stem Cells and Tissues. Angew. Chem. Int. Ed Engl. 56, 11268–11271. https://doi.org/10.1002/anie.201700424

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., Selbach, M., 2011. Global quantification of mammalian gene expression control. Nature 473, 337–342. https://doi.org/10.1038/nature10098

Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., Carr, S.A., Lander, E.S., Fink, G.R., Regev, A., 2013. High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell 155, 1409–1421. https://doi.org/10.1016/j.cell.2013.10.047

Sedgwick, B., Bates, P.A., Paik, J., Jacobs, S.C., Lindahl, T., 2007. Repair of alkylated DNA: recent advances. DNA Repair 6, 429–442. https://doi.org/10.1016/j.dnarep.2006.10.005

Sergiev, P.V., Serebryakova, M.V., Bogdanov, A.A., Dontsova, O.A., 2008. The ybiN gene of Escherichia coli encodes adenine-N6 methyltransferase specific for modification of A1618 of

23 S ribosomal RNA, a methylated residue located close to the ribosomal exit tunnel. J. Mol. Biol. 375, 291–300. https://doi.org/10.1016/j.jmb.2007.10.051
Shah, J.C., Clancy, M.J., 1992. IME4, a gene that mediates MAT and nutritional control of meiosis in Saccharomyces cerevisiae. Mol. Cell. Biol. 12, 1078–1086. https://doi.org/10.1128/mcb.12.3.1078
Shaw, N.N., Arya, D.P., 2008. Recognition of the unique structure of DNA:RNA hybrids. Biochimie 90, 1026–1039. https://doi.org/10.1016/j.biochi.2008.04.011
Shi, H., Zhang, X., Weng, Y.-L., Lu, Zongyang, Liu, Y., Lu, Zhike, Li, J., Hao, P., Zhang, Y., Zhang, F., Wu, Y., Delgado, J.Y., Su, Y., Patel, M.J., Cao, X., Shen, B., Huang, X., Ming, G.-L., Zhuang, X., Song, H., He, C., Zhou, T., 2018. m6A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature 563, 249–253. https://doi.org/10.1038/s41586-018-0666-1
Shi, X., Seluanov, A., Gorbunova, V., 2007. Cell divisions are required for L1 retrotransposition. Mol. Cell. Biol. 27, 1264–1270. https://doi.org/10.1128/MCB.01888-06
Shim, H., 2011. One target, different effects: a comparison of distinct therapeutic antibodies against the same targets. Exp. Mol. Med. 43, 539–549. https://doi.org/10.3858/emm.2011.43.10.063
Shimba, S., Bokar, J.A., Rottman, F., Reddy, R., 1995. Accurate and efficient N-6-adenosine methylation in spliceosomal U6 small nuclear RNA by HeLa cell extract in vitro. Nucleic Acids Res. 23, 2421–2426. https://doi.org/10.1093/nar/23.13.2421
Shinkura, R., Tian, M., Smith, M., Chua, K., Fujiwara, Y., Alt, F.W., 2003. The influence of transcriptional orientation on endogenous switch region function. Nat. Immunol. 4, 435–441. https://doi.org/10.1038/ni918
Skourti-Stathaki, K., Kamieniarz-Gdula, K., Proudfoot, N.J., 2014. R-loops induce repressive chromatin marks over mammalian gene terminators. Nature 516, 436–439. https://doi.org/10.1038/nature13787
Skourti-Stathaki, K., Proudfoot, N.J., 2014. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 28, 1384–1396. https://doi.org/10.1101/gad.242990.114
Skourti-Stathaki, K., Proudfoot, N.J., Gromak, N., 2011. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell 42, 794–805. https://doi.org/10.1016/j.molcel.2011.04.026
Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., Cimprich, K.A., 2014. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. Mol. Cell 56, 777–785. https://doi.org/10.1016/j.molcel.2014.10.020
Sommer, S., Lavi, U., Darnell, J.E., 1978. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. J. Mol. Biol. 124, 487–499. https://doi.org/10.1016/0022-2836(78)90183-3
Sood, A.J., Viner, C., Hoffman, M.M., 2019. DNAmod: the DNA modification database. J. Cheminformatics 11, 30. https://doi.org/10.1186/s13321-019-0349-4
Sorrells, S., Nik, S., Casey, M.J., Cameron, R.C., Truong, H., Toruno, C., Gulfo, M., Lowe, A., Jette, C., Stewart, R.A., Bowman, T.V., 2018. Spliceosomal components protect embryonic neurons from R-loop-mediated DNA damage and apoptosis. Dis. Model. Mech. 11. https://doi.org/10.1242/dmm.031583
Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.-W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., Chang, H.Y., 2015. Structural imprints in vivo decode RNA regulatory mechanisms. Nature 519, 486–490. https://doi.org/10.1038/nature14263
Srinivasan, P.R., Borek, E., 1964. ENZYMATIC ALTERATION OF NUCLEIC ACID STRUCTURE. Science 145, 548–553. https://doi.org/10.1126/science.145.3632.548
Stavnezer, J., Guikema, J.E.J., Schrader, C.E., 2008. Mechanism and Regulation of Class Switch Recombination. Annu. Rev. Immunol. 26, 261–292. https://doi.org/10.1146/annurev.immunol.26.021607.090248

Stöger, R., 2017. N6-methyldeoxyadenine (6mA) Is a Rare Beast - in Animals at Least (retrospective on DOI 10.1002/bies.201500076). BioEssays News Rev. Mol. Cell. Dev. Biol. 39. https://doi.org/10.1002/bies.201700157

Stork, C.T., Bocek, M., Crossley, M.P., Sollier, J., Sanz, L.A., Chédin, F., Swigut, T., Cimprich, K.A., 2016. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. eLife 5. https://doi.org/10.7554/eLife.17548

Sugino, A., Hirose, S., Okazaki, R., 1972. RNA-linked nascent DNA fragments in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 69, 1863–1867. https://doi.org/10.1073/pnas.69.7.1863

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., Dean, C., 2013. R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. Science 340, 619–621. https://doi.org/10.1126/science.1234848

Sun, Q., Huang, S., Wang, X., Zhu, Y., Chen, Z., Chen, D., 2015. N6-methyladenine functions as a potential epigenetic mark in eukaryotes. BioEssays News Rev. Mol. Cell. Dev. Biol. 37, 1155–1162. https://doi.org/10.1002/bies.201500076

Sundquist, W.I., Klug, A., 1989. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature 342, 825–829. https://doi.org/10.1038/342825a0

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., Rao, A., 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324, 930–935. https://doi.org/10.1126/science.1170116

Tam, A.S., Stirling, P.C., 2019. Splicing, genome stability and disease: splice like your genome depends on it! Curr. Genet. 65, 905–912. https://doi.org/10.1007/s00294-019-00964-0

Tan-Wong, S.M., Dhir, S., Proudfoot, N.J., 2019. R-Loops Promote Antisense Transcription across the Mammalian Genome. Mol. Cell 76, 600-616.e6. https://doi.org/10.1016/j.molcel.2019.10.002

Thomas, M., White, R.L., Davis, R.W., 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. Proc. Natl. Acad. Sci. U. S. A. 73, 2294–2298. https://doi.org/10.1073/pnas.73.7.2294

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., De Vos, J., Thomas, A., Theillet, C., Pommier, Y., Tazi, J., Coquelle, A., Pasero, P., 2009. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol. 11, 1315–1324. https://doi.org/10.1038/ncb1984

Unger, G., Venner, H., 1966. [Remarks on minor bases in spermatic desoxyribonucleic acid]. Hoppe. Seylers Z. Physiol. Chem. 344, 280–283.

Vanoosthuyse, V., 2018. Strengths and Weaknesses of the Current Strategies to Map and Characterize R-Loops. Non-Coding RNA 4. https://doi.org/10.3390/ncrna4020009

Vanyushin, B.F., Tkacheva, S.G., Belozersky, A.N., 1970. Rare bases in animal DNA. Nature 225, 948–949. https://doi.org/10.1038/225948a0

Verma, N., Pan, H., Doré, L.C., Shukla, A., Li, Q.V., Pelham-Webb, B., Teijeiro, V., González, F., Krivtsov, A., Chang, C.-J., Papapetrou, E.P., He, C., Elemento, O., Huangfu, D., 2018. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. Nat. Genet. 50, 83–95. https://doi.org/10.1038/s41588-017-0002-y

Vu, L.P., Pickering, B.F., Cheng, Y., Zaccara, S., Nguyen, D., Minuesa, G., Chou, T., Chow, A., Saletore, Y., MacKay, M., Schulman, J., Famulare, C., Patel, M., Klimek, V.M., Garrett-Bakelman, F.E., Melnick, A., Carroll, M., Mason, C.E., Jaffrey, S.R., Kharas, M.G., 2017. The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nat. Med. 23, 1369–1376. https://doi.org/10.1038/nm.4416

Waddington, C.H., 1959. Canalization of development and genetic assimilation of acquired characters. Nature 183, 1654–1655. https://doi.org/10.1038/1831654a0

Wahba, L., Amon, J.D., Koshland, D., Vuica-Ross, M., 2011. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol. Cell 44, 978–988. https://doi.org/10.1016/j.molcel.2011.10.017

Wahba, L., Costantino, L., Tan, F.J., Zimmer, A., Koshland, D., 2016. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. Genes Dev. 30, 1327–1338. https://doi.org/10.1101/gad.280834.116

Wahba, L., Gore, S.K., Koshland, D., 2013. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. eLife 2, e00505. https://doi.org/10.7554/eLife.00505

Walsh, C.P., Chaillet, J.R., Bestor, T.H., 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. 20, 116–117. https://doi.org/10.1038/2413

Wan, Y., Zheng, X., Chen, H., Guo, Y., Jiang, H., He, X., Zhu, X., Zheng, Y., 2015. Splicing function of mitotic regulators links R-loop–mediated DNA damage to tumor cell killing. J. Cell Biol. 209, 235–246. https://doi.org/10.1083/jcb.201409073

Wang, I.X., Grunseich, C., Fox, J., Burdick, J., Zhu, Z., Ravazian, N., Hafner, M., Cheung, V.G., 2018. Human proteins that interact with RNA/DNA hybrids. Genome Res. 28, 1405–1414. https://doi.org/10.1101/gr.237362.118

Wang, J., Huang, J., Shi, G., 2020. Retrotransposons in pluripotent stem cells. Cell Regen. 9. https://doi.org/10.1186/s13619-020-00046-4

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., He, C., 2014. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120. https://doi.org/10.1038/nature12730

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., He, C., 2015. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388–1399. https://doi.org/10.1016/j.cell.2015.05.014

Warda, A.S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Höbartner, C., Sloan, K.E., Bohnsack, M.T., 2017. Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. 18, 2004–2014. https://doi.org/10.15252/embr.201744940

Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737–738. https://doi.org/10.1038/171737a0

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., Schübeler, D., 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. 37, 853–862. https://doi.org/10.1038/ng1598

Wei, C.M., Gershowitz, A., Moss, B., 1975. Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. Cell 4, 379–386. https://doi.org/10.1016/0092-8674(75)90158-0

Wellinger, R.E., Prado, F., Aguilera, A., 2006. Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. Mol. Cell. Biol. 26, 3327–3334. https://doi.org/10.1128/MCB.26.8.3327-3334.2006

Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B.S., Dong, L., Shi, H., Skibbe, J., Shen, C., Hu, C., Sheng, Y., Wang, Y., Wunderlich, M., Zhang, B., Dore, L.C., Su, R., Deng, X., Ferchen, K., Li, C., Sun, M., Lu, Z., Jiang, X., Marcucci, G., Mulloy, J.C., Yang, J., Qian, Z., Wei, M., He, C., Chen, J., 2018. METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m6A Modification. Cell Stem Cell 22, 191-205.e9. https://doi.org/10.1016/j.stem.2017.11.016

Westover, K.D., Bushnell, D.A., Kornberg, R.D., 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science 303, 1014–1016. https://doi.org/10.1126/science.1090839

Wheldon, L.M., Abakir, A., Ferjentsik, Z., Dudnakova, T., Strohbuecker, S., Christie, D., Dai, N., Guan, S., Foster, J.M., Corrêa, I.R., Loose, M., Dixon, J.E., Sottile, V., Johnson, A.D., Ruzov, A., 2014. Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. Cell Rep. 7, 1353–1361. https://doi.org/10.1016/j.celrep.2014.05.003

White, R.L., Hogness, D.S., 1977. R loop mapping of the 18S and 28S sequences in the long and short repeating units of Drosophila melanogaster rDNA. Cell 10, 177–192. https://doi.org/10.1016/0092-8674(77)90213-6

Wilson, V.L., Jones, P.A., 1983. DNA methylation decreases in aging but not in immortal cells. Science 220, 1055–1057. https://doi.org/10.1126/science.6844925

Wu, R., Li, A., Sun, B., Sun, J.-G., Zhang, J., Zhang, T., Chen, Y., Xiao, Y., Gao, Y., Zhang, Q., Ma, J., Yang, X., Liao, Y., Lai, W.-Y., Qi, X., Wang, S., Shu, Y., Wang, H.-L., Wang, F., Yang, Y.-G., Yuan, Z., 2019. A novel m6A reader Prrc2a controls oligodendroglial specification and myelination. Cell Res. 29, 23–41. https://doi.org/10.1038/s41422-018-0113-8

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., Tackett, A., Wang, G., Hon, L.S., Fang, G., Swenberg, J.A., Xiao, A.Z., 2016a. DNA methylation on N(6)-adenine in mammalian embryonic stem cells. Nature 532, 329–333. https://doi.org/10.1038/nature17640

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., Tackett, A., Wang, G., Hon, L.S., Fang, G., Swenberg, J.A., Xiao, A.Z., 2016b. DNA methylation on N(6)-adenine in mammalian embryonic stem cells. Nature 532, 329–333. https://doi.org/10.1038/nature17640

Xiang, Y., Laurent, B., Hsu, C.-H., Nachtergaele, S., Lu, Z., Sheng, W., Xu, C., Chen, H., Ouyang, J., Wang, S., Ling, D., Hsu, P.-H., Zou, L., Jambhekar, A., He, C., Shi, Y., 2017. RNA m6A methylation regulates the ultraviolet-induced DNA damage response. Nature 543, 573–576. https://doi.org/10.1038/nature21671

Xiao, C.-L., Zhu, S., He, M., Chen, D., Zhang, Q., Chen, Y., Yu, G., Liu, J., Xie, S.-Q., Luo, F., Liang, Z., Wang, D.-P., Bo, X.-C., Gu, X.-F., Wang, K., Yan, G.-R., 2018. N6-Methyladenine DNA Modification in the Human Genome. Mol. Cell 71, 306-318.e7. https://doi.org/10.1016/j.molcel.2018.06.015

Xie, Y., Mates, L., Ivics, Z., Izsvák, Z., Martin, S.L., An, W., 2013. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. Mob. DNA 4, 10. https://doi.org/10.1186/1759-8753-4-10

Xu, B., Clayton, D.A., 1995. A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. Mol. Cell. Biol. 15, 580–589. https://doi.org/10.1128/mcb.15.1.580

Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., Min, J., 2015. Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. J. Biol. Chem. 290, 24902–24913. https://doi.org/10.1074/jbc.M115.680389

Xu, J., Chmela, V., Green, N.J., Russell, D.A., Janicki, M.J., Góra, R.W., Szabla, R., Bond, A.D., Sutherland, J.D., 2020. Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides. Nature 582, 60–66. https://doi.org/10.1038/s41586-020-2330-9

Xu, K., Yang, Y., Feng, G.-H., Sun, B.-F., Chen, J.-Q., Li, Y.-F., Chen, Y.-S., Zhang, X.-X., Wang, C.-X., Jiang, L.-Y., Liu, C., Zhang, Z.-Y., Wang, X.-J., Zhou, Q., Yang, Y.-G., Li, W., 2017. Mettl3-mediated m6A regulates spermatogonial differentiation and meiosis initiation. Cell Res. 27, 1100–1114. https://doi.org/10.1038/cr.2017.100

Xu, W., Li, K., Li, S., Hou, Q., Zhang, Y., Liu, K., Sun, Q., 2020. The R-Loop Atlas of Arabidopsis Development and Responses to Environmental Stimuli. Plant Cell 32, 888. https://doi.org/10.1105/tpc.19.00802

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, Rong, Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, Ruiqiang, Qiao, J., Tang, F., 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. 20, 1131–1139. https://doi.org/10.1038/nsmb.2660

Yang, S., Wei, J., Cui, Y.-H., Park, G., Shah, P., Deng, Y., Aplin, A.E., Lu, Z., Hwang, S., He, C., He, Y.-Y., 2019. m6A mRNA demethylase FTO regulates melanoma tumorigenicity and response to anti-PD-1 blockade. Nat. Commun. 10, 2782. https://doi.org/10.1038/s41467-019-10669-0

Yang, X., Liu, Q.-L., Xu, W., Zhang, Y.-C., Yang, Y., Ju, L.-F., Chen, J., Chen, Y.-S., Li, K., Ren, J., Sun, Q., Yang, Y.-G., 2019. m6A promotes R-loop formation to facilitate transcription termination. Cell Res. 29, 1035–1038. https://doi.org/10.1038/s41422-019-0235-7

Yao, B., Cheng, Y., Wang, Z., Li, Y., Chen, L., Huang, L., Zhang, W., Chen, D., Wu, H., Tang, B., Jin, P., 2017. DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress. Nat. Commun. 8, 1122. https://doi.org/10.1038/s41467-017-01195-y

Yap, K.L., Zhou, M.-M., 2011. Structure and mechanisms of lysine methylation recognition by the chromodomain in gene transcription. Biochemistry 50, 1966–1980. https://doi.org/10.1021/bi101885m

Yasuhara, T., Kato, R., Hagiwara, Y., Shiotani, B., Yamauchi, M., Nakada, S., Shibata, A., Miyagawa, K., 2018. Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair. Cell 175, 558-570.e11. https://doi.org/10.1016/j.cell.2018.08.056

Yoon, K.-J., Ringeling, F.R., Vissers, C., Jacob, F., Pokrass, M., Jimenez-Cyrus, D., Su, Y., Kim, N.-S., Zhu, Y., Zheng, L., Kim, S., Wang, X., Doré, L.C., Jin, P., Regot, S., Zhuang, X., Canzar, S., He, C., Ming, G.-L., Song, H., 2017. Temporal Control of Mammalian Cortical Neurogenesis by m6A Methylation. Cell 171, 877-889.e17. https://doi.org/10.1016/j.cell.2017.09.003

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T.E., Lieber, M.R., 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat. Immunol. 4, 442–451. https://doi.org/10.1038/ni919

Yu, K., Roy, D., Huang, F.-T., Lieber, M.R., 2006. Detection and structural analysis of R-loops. Methods Enzymol. 409, 316–329. https://doi.org/10.1016/S0076-6879(05)09018-X

Yu, M., Ji, L., Neumann, D.A., Chung, D.-H., Groom, J., Westpheling, J., He, C., Schmitz, R.J., 2015. Base-resolution detection of N4-methylcytosine in genomic DNA using 4mC-Tet-assisted-bisulfite- sequencing. Nucleic Acids Res. 43, e148. https://doi.org/10.1093/nar/gkv738

Yue, Y., Liu, J., He, C., 2015. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. Genes Dev. 29, 1343–1355. https://doi.org/10.1101/gad.262766.115

Zaccara, S., Jaffrey, S.R., 2020. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. Cell 0. https://doi.org/10.1016/j.cell.2020.05.012

Zarnack, K., König, J., Tajnik, M., Martincorena, I., Eustermann, S., Stévant, I., Reyes, A., Anders, S., Luscombe, N.M., Ule, J., 2013. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell 152, 453–466. https://doi.org/10.1016/j.cell.2012.12.023

Zatreanu, D., Han, Z., Mitter, R., Tumini, E., Williams, H., Gregersen, L., Dirac-Svejstrup, A.B., Roma, S., Stewart, A., Aguilera, A., Svejstrup, J.Q., 2019. Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. Mol. Cell 76, 57-69.e9. https://doi.org/10.1016/j.molcel.2019.07.037

Zeller, P., Padeken, J., van Schendel, R., Kalck, V., Tijsterman, M., Gasser, S.M., 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. Nat. Genet. 48, 1385–1395. https://doi.org/10.1038/ng.3672

Zeng, M., Dai, X., Liang, Z., Sun, R., Huang, S., Luo, L., Li, Z., 2020. Critical roles of mRNA m6A modification and YTHDC2 expression for meiotic initiation and progression in female germ cells. Gene 753, 144810. https://doi.org/10.1016/j.gene.2020.144810

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., To, J.K., Laird, M.R., Lavidas, I., Liu, Z., Loveland, J.E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D.N., Newman, V., Nuhn, M., Ogeh, D., Ong, C.K., Parker, A., Patricio, M., Riat, H.S., Schuilenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S.E., Kostadima, M., Langridge, N., Martin, F.J., Muffato, M., Perry, E., Ruffier, M., Staines, D.M., Trevanion, S.J., Aken, B.L., Cunningham, F., Yates, A., Flicek, P., 2018. Ensembl 2018. Nucleic Acids Res. 46, D754–D761. https://doi.org/10.1093/nar/gkx1098

Zerbino, D.R., Wilder, S.P., Johnson, N., Juettemann, T., Flicek, P.R., 2015. The ensembl regulatory build. Genome Biol. 16, 56. https://doi.org/10.1186/s13059-015-0621-5

Zhang, C., Chen, Liping, Peng, D., Jiang, A., He, Y., Zeng, Y., Xie, C., Zhou, H., Luo, X., Liu, H., Chen, Liang, Ren, J., Wang, W., Zhao, Y., 2020. METTL3 and N6-Methyladenosine Promote Homologous Recombination-Mediated Repair of DSBs by Modulating DNA-RNA Hybrid Accumulation. Mol. Cell S1097276520304020. https://doi.org/10.1016/j.molcel.2020.06.017

Zhang, G., Huang, H., Liu, D., Cheng, Y., Liu, X., Zhang, W., Yin, R., Zhang, D., Zhang, P., Liu, J., Li, C., Liu, B., Luo, Y., Zhu, Y., Zhang, N., He, S., He, C., Wang, H., Chen, D., 2015. N6-methyladenine DNA modification in Drosophila. Cell 161, 893–906. https://doi.org/10.1016/j.cell.2015.04.018

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137. https://doi.org/10.1186/gb-2008-9-9-r137

Zhang, Y., Wang, X., Zhang, X., Wang, J., Ma, Y., Zhang, L., Cao, X., 2019. RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. Proc. Natl. Acad. Sci. U. S. A. 116, 976–981. https://doi.org/10.1073/pnas.1812536116

Zhang, Z., Theler, D., Kaminska, K.H., Hiller, M., de la Grange, P., Pudimat, R., Rafalska, I., Heinrich, B., Bujnicki, J.M., Allain, F.H.-T., Stamm, S., 2010. The YTH domain is a novel RNA binding domain. J. Biol. Chem. 285, 14701–14710. https://doi.org/10.1074/jbc.M110.104711

Zhao, B.S., Wang, X., Beadell, A.V., Lu, Z., Shi, H., Kuuspalu, A., Ho, R.K., He, C., 2017. m6A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. Nature 542, 475–478. https://doi.org/10.1038/nature21355

Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., Hao, Y.-J., Ping, X.-L., Chen, Y.-S., Wang, W.-J., Jin, K.-X., Wang, X., Huang, C.-M., Fu, Y., Ge, X.-M., Song, S.-H., Jeong, H.S., Yanagisawa, H., Niu, Y., Jia, G.-F., Wu, W., Tong, W.-M., Okamoto, A., He, C., Rendtlew Danielsen, J.M., Wang, X.-J., Yang, Y.-G., 2014. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res. 24, 1403–1419. https://doi.org/10.1038/cr.2014.151

Zhao, Y., Shi, Y., Shen, H., Xie, W., 2020. m6A-binding proteins: the emerging crucial performers in epigenetics. J. Hematol. Oncol.J Hematol Oncol 13, 35. https://doi.org/10.1186/s13045-020-00872-8

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C.J., Vågbø, C.B., Shi, Y., Wang, W.-L., Song, S.-H., Lu, Z., Bosmans, R.P.G., Dai, Q., Hao, Y.-J., Yang, X., Zhao, W.-M., Tong, W.-M., Wang, X.-J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H.E., Klungland, A., Yang, Y.-G., He, C., 2013. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol. Cell 49, 18–29. https://doi.org/10.1016/j.molcel.2012.10.015

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., Fray, R.G., 2008. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. Plant Cell 20, 1278–1288. https://doi.org/10.1105/tpc.108.058883

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., Qian, S.-B., 2015. Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature 526, 591–594. https://doi.org/10.1038/nature15377

Zhu, Quan, Hoong, N., Aslanian, A., Hara, T., Benner, C., Heinz, S., Miga, K.H., Ke, E., Verma, S., Soroczynski, J., Yates, J.R., Hunter, T., Verma, I.M., 2018. Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. Mol. Cell 70, 842-853.e7. https://doi.org/10.1016/j.molcel.2018.04.023

Zhu, Qifan, Stöger, R., Alberio, R., 2018. A Lexicon of DNA Modifications: Their Roles in Embryo Development and the Germline. Front. Cell Dev. Biol. 6, 24. https://doi.org/10.3389/fcell.2018.00024

Zhu, W., Wang, J.-Z., Xu, Z., Cao, M., Hu, Q., Pan, C., Guo, M., Wei, J.-F., Yang, H., 2019. Detection of N6-methyladenosine modification residues (Review). Int. J. Mol. Med. 43, 2267–2278. https://doi.org/10.3892/ijmm.2019.4169

Zhu, Z.-M., Huo, F.-C., Pei, D.-S., 2020. Function and evolution of RNA N6-methyladenosine modification. Int. J. Biol. Sci. 16, 1929–1940. https://doi.org/10.7150/ijbs.45231

Zimmer, A.D., Koshland, D., 2016. Differential roles of the RNases H in preventing chromosome instability. Proc. Natl. Acad. Sci. U. S. A. 113, 12220–12225. https://doi.org/10.1073/pnas.1613448113

Zuckerman, B., Ron, M., Mikl, M., Segal, E., Ulitsky, I., 2020. Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. Mol. Cell. https://doi.org/10.1016/j.molcel.2020.05.013