

# **RNA Methylation During Epigenetic Reprogramming**

**Mr. Karl David Hudson**

**Imperial College London**

**Institute of Clinical Sciences**

**MRC London Institute of Medical Sciences**

**This document is submitted for the degree of  
Philosophical Doctorate**

**Imperial College  
London**



## Declaration of originality

I hereby declare that, unless otherwise stated, the work presented within this document is my own. Where any work or information is not my own, it is clearly and appropriately referenced to indicate the original source.

## Copyright declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

## Acknowledgements and dedication

I dedicate this body of work to my young family, having grown +3 -1 over the course of my PhD. Especially so for Wren, with us for such a short time. To my loving wife Holly, you have been a constant support. To baby Noah, I hope that this will help you grow to be proud of your dad. I can think of no more fitting tribute to the past four years.

This thesis could never have come about without the guidance and support of my supervisor Prof. Petra Hajkova. My mentors, Prof. Richard Festenstein, Dr. Till Bartke, and Dr. Peter Sarkies have also given constructive feedback.

I should thank all overlapping members of the Reprogramming and Chromatin group, both past and present, in addition to members of the closely-linked Germline and Pluripotency group at the MRC LMS. Of particular note, our lab managers Dr. Catherine Walker and Dr. Eric Vazquez-Ferrer have been dutiful in ensuring the smooth operation of the laboratory. Dr. Rachael Amouroux, Dr. Cristina Requena-Torres, and Dr. Ammar Nasif should be thanked for varying degrees of mass spectrometry guidance. Dr. Harry Leitch has been a veritable font of wisdom on stem cell culture and the peculiarities of genetic background, while Dr. Malgorzata (Gosia) Borkowska and Dr. Tien-Chi Huang taught me embryo dissection techniques. The Central Biomedical Services staff have been a constant assistance in animal husbandry and I thank them for their care. Where any work has been performed by a collaborative colleague it is clearly outlined as such.

Thank you to Petra and Cristina for giving me feedback through the process of drafting this thesis. Additionally, I would like to thank my wife Holly, Dr. Liam Shaw, Dr. Claire Hall, and Niket Shah for proofreading, giving feedback, and providing design assistance. I feel truly lucky to have had such a support network.

As I conclude in the lab I wish each group member well and hope, and anticipate, great things to continue to come from the Hajkova lab.

## Abstract

Biologically derived, chemically-modified RNA molecules represent a ubiquitously present but poorly understood area of molecular biology. Despite this substantive gap in our understanding RNA methyltransferase (RNMT) enzymes, which catalyse the formation of RNA 5mC, are incredibly highly conserved, while disrupting m6A metabolism leads to early embryo lethality. This indicates that these modifications perform critical, but poorly characterised, biological functions.

Knockout experiments and clinical pathologies guided the work of this thesis in which RNA methylation was analysed in the context of epigenetic reprogramming and development more broadly. LC-MS/MS-based quantification techniques were developed to achieve unparalleled sensitivity and applied to *in vivo* and *in vitro* murine models of epigenetic reprogramming and developmental progression through pluripotency. While an intrinsic link between DNA 5mdC methylation and RNA 5mC methylation was refuted, 5mC changes dynamically along a naive-primed pluripotent axis, in line with the redistribution of NOP2 and NSUN2 within the cell. While these RNMTs are strongly linked in bringing about these changes in 5mC abundance, increased expression of NSUN2 in the trophectoderm at the blastocyst stage uncovers a potential lineage specific role for this enzyme.

This study reveals hitherto undiscovered connections between RNA methylation and the developmental transitions into and out of pluripotency. Better understanding of these molecular processes may ultimately lead into more refined and effective approaches to regenerative medicine.

# Table of contents

|                                                                                                                     |    |
|---------------------------------------------------------------------------------------------------------------------|----|
| Declaration of originality .....                                                                                    | 2  |
| Copyright declaration .....                                                                                         | 3  |
| Acknowledgements and dedication.....                                                                                | 4  |
| Abstract.....                                                                                                       | 6  |
| Table of contents .....                                                                                             | 7  |
| Frequently used abbreviations .....                                                                                 | 11 |
| List of figures.....                                                                                                | 14 |
| List of tables.....                                                                                                 | 19 |
| Chapter 1: Introduction & literature review .....                                                                   | 20 |
| 1.1    RNA modifications.....                                                                                       | 21 |
| 1.1.1 Historical perspectives on RNA modifications .....                                                            | 21 |
| 1.1.2 On the origin of RNA methyltransferases (RNMTs) – convergent activity of DNA and RNA methyltransferases ..... | 24 |
| 1.1.3 The biological significance of the RNA methyltransferase enzymes .....                                        | 26 |
| 1.1.4 RNA 5mC characterisation .....                                                                                | 30 |
| 1.1.5 RNA 5hmC characterisation .....                                                                               | 36 |
| 1.1.6 RNA m6A writers, readers, and erasers: The molecular machinery used to regulate differentiation.....          | 40 |
| 1.1.7    RNA m6A is required for early embryo development.....                                                      | 48 |
| 1.2 Pluripotency and epigenetic reprogramming .....                                                                 | 52 |
| 1.2.1 Defining pluripotency .....                                                                                   | 52 |
| 1.2.2 The epigenetic ground state .....                                                                             | 56 |
| 1.2.3 Mechanisms of epigenetic reprogramming <i>in vivo</i> .....                                                   | 59 |
| 1.2.4 Modelling epigenetic reprogramming <i>in vitro</i> .....                                                      | 67 |
| 1.3 Key Hypotheses.....                                                                                             | 68 |
| Chapter 2: Materials and methods.....                                                                               | 69 |
| 2.1 Mice .....                                                                                                      | 70 |
| 2.1.1 Animal husbandry .....                                                                                        | 70 |
| 2.1.2 Genotyping.....                                                                                               | 70 |
| 2.1.3 Dissections .....                                                                                             | 72 |
| 2.2 Cell culture .....                                                                                              | 72 |
| 2.2.1 Embryonic stem cells .....                                                                                    | 72 |

|                                                                      |    |
|----------------------------------------------------------------------|----|
| 2.2.2 Epiblast stem cells.....                                       | 74 |
| 2.2.3 <i>In vitro</i> fertilisation and early embryo culture.....    | 74 |
| 2.3 Florescence-activated cell sorting (FACS) .....                  | 75 |
| 2.4 Western blotting.....                                            | 75 |
| 2.4.1 Protein extraction and quantification.....                     | 75 |
| 2.4.2 Casting and running gel .....                                  | 76 |
| 2.5 Immunostaining .....                                             | 79 |
| 2.5.1 Embryonic stem cells .....                                     | 79 |
| 2.5.2 Genital Ridges .....                                           | 79 |
| 2.5.3 PGC single cells .....                                         | 80 |
| 2.5.4 Early embryo .....                                             | 80 |
| 2.6 RNA extraction .....                                             | 81 |
| 2.6.1 QIAGEN RNeasy kit.....                                         | 81 |
| 2.6.2 Zymo Research Duet DNA/RNA MiniPrep kit.....                   | 81 |
| 2.6.3 TRIzol.....                                                    | 81 |
| 2.7 RNA quantification and quality assessment .....                  | 82 |
| 2.7.1 Bioanalyzer.....                                               | 82 |
| 2.7.2 RNA gel electrophoresis.....                                   | 82 |
| 2.8 Nucleotide digestion.....                                        | 82 |
| 2.8.1 DNase treatment .....                                          | 82 |
| 2.8.2 RNase treatment.....                                           | 83 |
| 2.8.3 RNA clean up.....                                              | 83 |
| 2.9 RNA dot blot.....                                                | 83 |
| 2.9.1 Urea and formamide/formaldehyde methods .....                  | 83 |
| 2.9.2 NaOH method .....                                              | 84 |
| 2.10 RNA subtype selection .....                                     | 84 |
| 2.10.1 Poly(A) enrichment .....                                      | 84 |
| 2.10.2 Small RNA enrichment .....                                    | 84 |
| 2.11 Liquid Chromatography tandem Mass Spectrometry (LC-MS/MS) ..... | 85 |
| 2.11.1 Method development and optimisation.....                      | 85 |
| 2.11.2 Sample preparation .....                                      | 88 |
| 2.12 Real-time quantitative PCR (RT-qPCR) .....                      | 90 |
| 2.13 Bacterial work and <i>in vitro</i> transcription .....          | 93 |
| 2.13.1 Plasmid extraction .....                                      | 93 |

|                                                                                                                                              |     |
|----------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 2.13.2 Plasmid sequencing .....                                                                                                              | 93  |
| 2.13.3 <i>In vitro</i> transcription .....                                                                                                   | 93  |
| Chapter 3: Development of methods and tools .....                                                                                            | 94  |
| 3.1 Introduction .....                                                                                                                       | 95  |
| 3.1.1 Dot blotting.....                                                                                                                      | 95  |
| 3.1.2 Liquid chromatography tandem mass spectrometry .....                                                                                   | 96  |
| 3.2 Experimental aims .....                                                                                                                  | 101 |
| 3.3 Results.....                                                                                                                             | 101 |
| 3.3.1 Dot blotting development as a semi-quantitative measure of modified RNA nucleotide abundance .....                                     | 101 |
| 3.3.2 LC-MS/MS optimisation quantitative measure of modified RNA nucleotide abundance .....                                                  | 119 |
| 3.4 Discussion .....                                                                                                                         | 134 |
| Chapter 4: RNA modifications in stem cells.....                                                                                              | 135 |
| 4.1 Introduction .....                                                                                                                       | 136 |
| 4.2 Experimental Aims .....                                                                                                                  | 138 |
| 4.3 Results.....                                                                                                                             | 139 |
| 4.3.1 Descriptions and derivations of cell lines .....                                                                                       | 139 |
| 4.3.2 RNMT gene expression .....                                                                                                             | 145 |
| 4.3.3 RNMT subcellular localisation reveals novel NOP2 and NSUN2 distributions.....                                                          | 152 |
| 4.3.4 RNA modification quantification.....                                                                                                   | 162 |
| 4.4 Discussion .....                                                                                                                         | 183 |
| Chapter 5: RNA modifications <i>in vivo</i> .....                                                                                            | 186 |
| 5.1 Introduction .....                                                                                                                       | 187 |
| 5.2 Experimental Aims .....                                                                                                                  | 188 |
| 5.3 Results.....                                                                                                                             | 188 |
| 5.3.1 Gene expression analysis reveals dynamic changes in RNMT and m6A-related gene expression in early embryo and PGC development. ....     | 188 |
| 5.3.2 RNMT protein levels and subcellular localisations in early embryo and PGC development ....                                             | 192 |
| 5.3.3 PGC total RNA modification abundances provide tentative indications of 5mC and m6A elevations predating epigenetic reprogramming ..... | 212 |
| 5.4 Discussion .....                                                                                                                         | 215 |
| Chapter 6: Discussions and conclusions .....                                                                                                 | 217 |
| 6.1 Technical developments .....                                                                                                             | 218 |
| 6.2 Quantification of modified RNAs .....                                                                                                    | 220 |
| 6.3 TET-independent formation of RNA 5hmC and 5fC.....                                                                                       | 222 |

|                                    |     |
|------------------------------------|-----|
| 6.4 Subcellular localisation ..... | 223 |
| Bibliography .....                 | 226 |
| Appendices.....                    | 252 |

## Frequently used abbreviations

2i : 2 Inhibitor-containing media

5hmC : 5-Hydroxymethylcytosine

5mC : 5-Methylcytosine

5fC : 5-Formylcytosine

5caC : 5-Carboxylcytosine

A : Adenosine

ALKBH : AlkB homolog

bFGF : Fibroblast growth factor-basic

BSA : Bovine serum albumin

C : Cytosine

DAPI : 4',6-diamidino-2-phenylindole

DEPC : Diethyl pyrocarbonate

DNMT : DNA methyltransferase

EDTA : Ethylenediaminetetraacetic acid

EpiSC : Epiblast stem cell

ESC : Embryonic stem cell

FACS : Fluorescence-activated cell sorting

FCS : Foetal calf serum

FTO : Fat mass and obesity-associated protein

GMEM : Glasgow minimum essential medium

LC-MS/MS : Liquid chromatography tandem mass spectrometry

LOD : Limit of detection

LOQ : Limit of quantification

Lif : Leukaemia inhibitory factor

lncRNA : Long non-coding RNA

m6A : N6-Methyladenosine

METTL : Methyltransferase-like

MRM : Multiple reaction monitoring

m/z : Mass to charge ratio

PBS : Phosphate-buffered saline

Poly(A) : Polyadenylated

PBS-T : Phosphate-buffered saline with Tween-20

PGC : Primordial germ cell

RIPA : Radioimmunoprecipitation assay

RNMT : RNA methyltransferase

SDS : Sodium dodecyl sulphate

TBE : Tris/borate/EDTA

TBS : Tris-buffered saline

TBS-T : Tris-buffered saline with Tween-20

TET : Ten-eleven translocation enzyme

UTR : Untranslated region

YTHDF : YT521-B homology domain family

YTHDC : YT521-B homology domain containing

## List of figures

### Chapter 1: Introduction and literature review:

- Fig. 1.1 RNMTs are the evolutionary precursors of DNMTs
- Fig. 1.2 RNMT-catalysed RNA 5-methylcytidine formation
- Fig. 1.3 Morphology of *Trdmt1* and *Nsun2* double knockout embryos
- Fig. 1.4 The reported roles of RNMTs in chromatin reorganisation and transcriptional upregulation
- Fig. 1.5 TET-mediated oxidations of RNA 5mC
- Fig 1.6 Reversible m6A methylation
- Fig 1.7 RNA m6A modifications groups transcripts for cell fate transitions
- Fig 1.8 YTHDF2 is required to clear maternal mRNAs and permit zygotic development beyond the 2-cell stage
- Fig. 1.9 The mRNA of core stem cell genes contain m6A
- Fig 1.10 Developmental progression of pluripotency in mouse and human embryos
- Fig 1.11 Waddington's Classical Epigenetic Landscape
- Fig. 1.12 A schematic of mouse pre-implantation and germ cell development

### Chapter 3: Development of methods and tools

- Fig 3.1 Eddy diffusion in a packed chromatography column.
- Fig 3.2 Quadrupole mass spectrometer. Diagram of instrument operation
- Figure 3.3 Trialling RNA dot blot denaturation methodology shows F&F to be optimal
- Fig 3.4 Validation of signal linearity in RNA 5mC and 5hmC dot blotting
- Fig 3.5 RNA 5mC dot blot signal is due to contaminating DNA 5mdC

- Fig 3.6 *In vitro* transcribed 5mC and 5hmC dot blotting with RNase A control, NaOH denaturation, and baking.
- Fig 3.7 *In vitro* transcribed RNA from an unmethylated plasmid gives a true RNA 5mC signal
- Fig 3.8 Optimisation of RNA crosslinking and definition of  $\alpha$ -5mC and  $\alpha$ -5hmC RNA dot blot LOD values
- Fig 3.9 Anti-5mC antibody trials reveal Abcam (RM231) to be the most sensitive antibody
- Fig 3.10 5mC and 5hmC are undetectable in mESCs
- Fig 3.11 Active Motif Abflex  $\alpha$ -5mC antibody trial and determination of LOD
- Fig 3.12 Functionally optimised  $\alpha$ -5mC RNA dot blot gives a greater signal in ESCs grown in 2i than serum media
- Figure 3.13 Electrospray ionisation total internal chromatograms of nucleoside standards run together and in isolation
- Fig 3.14 Optimisation of mass spectrometry electrospray ionisation and cation extraction
- Fig 3.15 Optimisation of mass spectrometry iFunnel voltages for cytidine
- Fig 3.16 Optimisation of mass spectrometry iFunnel voltages for 5-methylcytidine
- Fig 3.17 Optimisation of mass spectrometry iFunnel voltages for 5-hydroxymethylcytidine
- Fig 3.18 Optimisation of mass spectrometry iFunnel voltages for 5-formylcytidine
- Fig 3.19 Optimisation of mass spectrometry collision gas
- Fig 3.20 Chromatograms of J1 serum mESC RNA modifications

#### Chapter 4: RNA modifications in stem cells

- Fig 4.1 Brightfield images of TKO and wild type stem cells in 2i, serum or EpiSC conditions

- Fig 4.2 TKO EpiLSCs transition from a ESC fate to adopt a EpiSCs fate as determined by RT-qPCR
- Fig 4.3 TKO EpiLSCs initiate some programs indicating transition beyond the EpiSC fate
- Fig 4.4 Fig 4.4 RNMT mRNA expression level in different pluripotent stem cell fates as measured by microarray and RT-qPCR reveals subtitle variations in expression
- Fig 4.5 RNMT mRNA expression level upon ablation of DNA methylation as measured by RT-qPCR reveals increased expression of multiple RNMTs
- Fig 4.6 RNMT protein expression level as measured by western blotting
- Fig 4.7 NOP2 subcellular localisation
- Fig 4.8 NSUN2 subcellular localisation
- Fig 4.9 NSUN2 subcellular localisation – WT 2i signal boosted
- Fig 4.10 NSUN5 subcellular localisation
- Fig 4.11 NSUN5 subcellular localisation – WT 2i, WT serum, and TKO serum boosted signal
- Fig 4.12 TRDMT1 subcellular localisation
- Fig 4.13 Summation of stem cell RNMT subcellular localisations
- Fig 4.14 5-Methylcytosine quantification by LC-MS/MS
- Fig 4.15 5-Hydroxymethylcytosine quantification by LC-MS/MS
- Fig 4.16 5-Formylcytosine quantification by LC-MS/MS
- Fig 4.17 6-Methyladenosine quantification by LC-MS/MS
- Fig 4.18 5-Methylcytosine quantification in RNA subtypes by LC-MS/MS
- Fig 4.19 5-Hydroxymethylcytosine quantification in RNA subtypes by LC-MS/MS
- Fig 4.20 5-Formylcytosine quantification in RNA subtypes by LC-MS/MS

- Fig 4.21 6-Methyladenosine quantification in RNA subtypes by LC-MS/MS
- Fig 4.22 5-Methylcytosine quantification by LC-MS/MS in TET triple knockout ESCs
- Fig 4.23 5-Hydroxymethylcytosine quantification by LC-MS/MS in TET triple knockout ESCs
- Fig 4.24 5-Formylcytosine quantification by LC-MS/MS in TET triple knockout ESCs
- Fig 4.25 6-Methyladenosine quantification by LC-MS/MS in TET triple knockout ESCs
- Fig 4.26 Representative Bioanalyzer traces

## Chapter 5: RNA modifications *in vivo*

- Fig 5.1 PGC RNA-seq reveals dynamic expression changes of proteins involved in 5mC and m6A metabolism during epigenetic reprogramming
- Fig 5.2 Early embryo RNA-seq reveals dynamic expression changes of proteins involved in 5mC and m6A metabolism during epigenetic reprogramming
- Fig 5.3 NOP2 expression in female PGCs
- Fig 5.4 NOP2 expression in male PGCs
- Fig 5.5 NSUN2 expression in female PGCs
- Fig 5.6 NSUN2 expression in male PGCs
- Fig 5.7 NSUN5 expression in female PGCs
- Fig 5.8 NSUN5 expression in male PGCs
- Fig 5.9 TRDMT1 expression in male PGCs
- Fig 5.10 NOP2 subcellular localisation in female PGCs
- Fig 5.11 NOP2 subcellular localisation in male PGCs
- Fig 5.12 NSUN2 subcellular localisation in female PGCs
- Fig 5.13 NSUN2 subcellular localisation in male PGCs

- Fig 5.14 NSUN5 subcellular localisation in female PGCs
- Fig 5.15 NSUN5 subcellular localisation in male PGCs
- Fig 5.16 TRDMT1 subcellular localisation in male PGCs
- Fig 5.17 NOP2 subcellular localisation in early embryos
- Fig 5.18 NSUN2 subcellular localisation in early embryos
- Fig 5.19 NOP2 and NSUN2 expression in blastocysts
- Fig 5.20 Modified RNA quantification by LC-MS/MS in PGCs and surrounding somatic cells

## List of tables

### Chapter 1: Introduction and literature review

- Table 1.1 Putative functions of RNA modifications
- Table 1.2 Properties of mouse cells in different phases of pluripotency

### Chapter 2: Materials and methods

- Table 2.1 SRY PCR conditions
- Table 2.2 Stem cell lines
- Table 2.3 Antibodies used for dot blotting, western blotting, and immunostaining
- Table 2.4 UHPLC mobile phase composition
- Table 2.5 Mass spectrometry parameters
- Table 2.6 Nucleoside standard curve for LC-MS/MS analysis
- Table 2.7 RNA digestion master mix for LC-MS/MS
- Table 2.8 Primers used for RT-qPCR

### Chapter 3: Development of methods and tools

- Table 3.1 Nucleoside cation abundance, transitions and retention times
- Table 3.2 The MS/MS parameters optimised with their initial values
- Table 3.3 Optimised mass spectrometry parameters for 5hmC and 5mC detection
- Table 3.4 Limit of detection and limit of quantification for RNA nucleosides before and after mass spectrometry optimisation

### Appendices

- Appendix table 1 Permission notes for cited figures

# Chapter 1: Introduction & literature review

## 1.1 RNA modifications

### 1.1.1 Historical perspectives on RNA modifications

It has been known for over half a century that RNA molecules undergo post-transcriptional modifications (Fleissner et al., 1962). Along with their discovery *in vivo*, this early work by Fleissner et al. found RNA methylation to be the result of enzyme-catalysed reactions, indicating biological relevance (Fleissner et al., 1962). To date, more than 140 different RNA chemical modifications have been detected *in vivo* (Machnicka et al., 2013, Limbach et al., 1994). Despite this, very little is known about what functions each of these modifications may have, save for a few notable exemptions, such as methylation at the 7' position of guanosine (m7G) in the cap structure of mRNA. This modification enables the cap binding protein (CBP) to bind to and protect the mRNA from exoribonucleases, and thus is crucial to enable appreciable levels of mRNA stability (Furuichi et al., 1977, Zhang et al., 2015). This exemplifies the capacity for chemical modifications to RNA nucleotides to be integral to biological processes, as opposed to merely being passive marks.

One of the fundamental reasons behind the paucity of functional understanding of these varied post-transcriptional modifications has been the lack of sufficiently sensitive quantitative methods for detecting them. In order to fully characterise such modifications their molarity alone is not sufficient, and so nucleotide-resolution sequencing strategies have been both intellectually necessary, and historically lacking. The technical barriers have been falling one by one as instrumentation and applied techniques have advanced, leading to a resurgence in the RNA modification field. The summation of functional biochemical modifications to RNA has been termed the epitranscriptome. It may at first appear to be a relatively novel area of research, with notable advancement over the past decade. However, its origins of study begin in the cradle of molecular biology less than a decade after Watson and Crick's discovery of the DNA double

helix (Watson et al., 1953). It is therefore pertinent to consider where intellectual gaps exist due to fundamental, immutable science, and where their inaccessibility is a temporary reflection of the techniques of the day. Method development and refinement form a critical arm of the research contained within this thesis.

The functionality of DNA modifications is much better understood than those present in RNA, with methylation to the 5 position of the cytosine base (5mdC) used as an epigenetic mark denoting silenced regions of eukaryotic genomes (Beyersmann, 2000). Such 5mdC modifications are crucial for mammalian cells to stably adopt differentiated states, and as such their erasure is a defining aspect of cellular reprogramming in both the zygote and primordial germ line (Hill et al., 2014; Sheaffer et al., 2014, Messerschmidt et al., 2014). However, DNA methylation is not as ubiquitous as one may be tempted to think. There are numerous examples of multicellular eukaryotes which lack DNA methylation, with both the *de novo* and maintenance DNA methyltransferases (DNMTs) having been lost on a number of separate occasions during the evolution of the nematode phylum for example (Rošić et al., 2018). This phenomenon is exemplified for animal, plant and fungal lineages in figure 1.1A. In stark contrast there are no known organisms which lack RNA methyltransferases (RNMTs) (Motorin et al., 2010). RNMTs are a subgroup of methyltransferase enzymes which catalyse the formation of RNA 5-methylcytosine (5mC). 5-Methylcytosine is particularly intriguing, as it is the only modification, along with its oxidised derivatives, observed in both the DNA and RNA of higher eukaryotes (Fig 1.2) (Machnicka et al., 2013, Motorin et al., 2010). Elucidating the function of these highly conserved enzymes, and how they may be connected to DNA 5mdC, formed the genesis of this thesis.

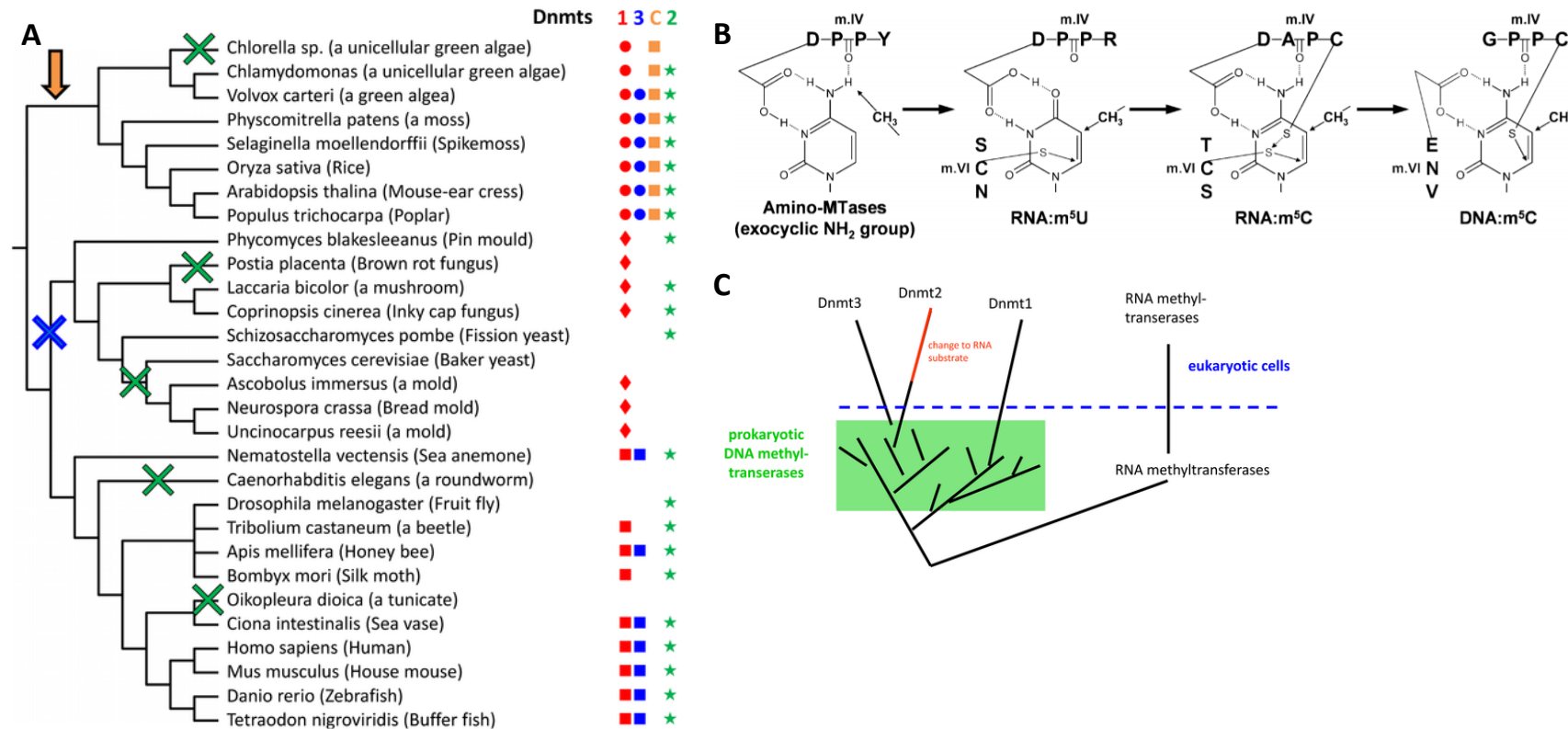


**Fig. 1.1 RNMTs are the evolutionary precursors of DNMTs.**

**(A)** Dnmts were categorized into the Dnmt1 (1, colored red), Dnmt3 (3, colored blue), Chromomethylase (C, colored orange) and Dnmt2 families (D, colored green). The Dnmt1 enzymes were subdivided into animal, plant and fungi subgroups, Dnmt3 enzymes into the canonical Dnmt3 enzymes and the plant DRM enzymes. Lineages that underwent loss of Dnmt2 are indicated by green crosses. Chromomethylases are related to Dnmt1 enzyme appearing in the plant lineage only (indicated by the orange arrow). Dnmt3 enzymes have been lost in the fungal lineage (indicated by a blue cross) (Jurkowski et al., 2011). Figure reproduced under a Creative Commons Attribution (CC-BY) Licence.

**(B)** Evolutionary model of relationships between pyrimidine-C5 MTases. The model of major functional innovations in the evolution of pyrimidine-C5 methylation. Key residues in motifs IV and VI are indicated. Thick lines indicate the hypothetical direction of evolution (Bujnicki et al., 2004). Figure reproduced with permission of the rights holder, Oxford University Press.

**(C)** Consensus model of the phylogeny of DNA methyltransferases and Dnmt2 (Jurkowski et al., 2011). Figure reproduced under a Creative Commons Attribution (CC-BY) Licence.

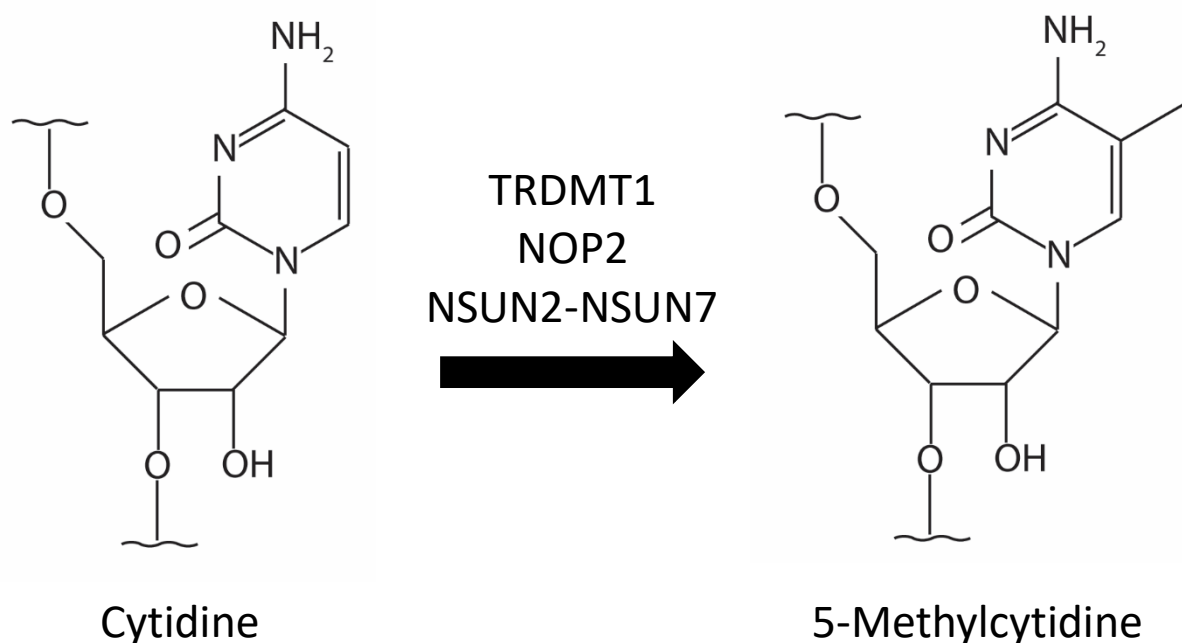


**Fig. 1.2 RNMT-catalysed RNA 5-methylcytidine formation.**

Mammalian 5mC-specific RNMTs are shown catalysing the formation of 5-methylcytidine through skeletal formulae. Wiggly lines indicate the extended ribonucleotide backbone.

### 1.1.2 On the origin of RNA methyltransferases (RNMTs) – convergent activity of DNA and RNA methyltransferases

To date DNA and RNA methyltransferases have been considered to be largely functionally distinct in terms of their targets along the boundaries indicated by their nomenclature. Analyses of structural homology and conserved motifs between these two families of SAM-dependent methyltransferases implicate evolutionary divergence of DNMTs from RNMTs of the Rossmann-fold-like superfamily (Bujnicki et al., 2004) (Fig. 1.1B). Furthermore, a combination of phylogenetic and biochemical analyses indicates that the RNMT DNMT2/ TRDMT1 (DNA methyltransferase 2/ tRNA aspartic acid methyltransferase 1), originally thought to be a DNMT, arose from a DNMT that underwent a change in substrate specificity such that the enzyme methylates tRNA but has minimal activity towards DNA (Jurkowski et al., 2008, Jurkowski et al.,

2011, Goll et al., 2006) (Fig. 1.1C). TRDMT1 is required to maintain the stability of its substrate tRNAs to prevent stress-induced cleavage, revealing a biological function as a RNA methyltransferase (Schaefer et al., 2010). Residual DNMT2/ TRDMT1 DNA methylation activity may remain, and so DNMT2/ TRDMT1 could have the capacity to methylate cytosines in either polynucleotide (Phalke et al, 2009). Indeed, knockout of *Trdmt1* in *Drosophila* reduces DNA methylation levels, with the mobility of retrotransposons enhanced. This is consistent not only with TRDMT1 methylating both DNA and RNA, but with methylation of each target showing biological significance (Phalke et al, 2009).

The canonical maintenance methyltransferase DNMT1 acts to methylate hemimethylated CpGs to enable the faithful transfer of 5mdC during DNA synthesis. Previous experiments have shown that DNMT1 can not only bind to all 3 possible double-stranded polynucleotide chains (RNA-RNA, RNA-DNA and DNA-DNA), but also displays methyltransferase activity towards all three substrates (P. Hajkova, personal communication). Furthermore, the RNMT NSUN2 has been shown to methylate hemi-methylated DNA *in vitro* (Sakita-Suto et al., 2007). The distinctions between 5mC, 5mdC and their respective deposition machineries may thus be more blurred than previously appreciated. In such a scenario, it is conceivable that there may be some form of functional interplay between RNA 5mC and DNA 5mdC. The global reduction of DNA methylation during epigenetic reprogramming may lead to a loosening of silencing of repressed regions of the genome. Could a potential link between methylation of DNA and RNA allow the repressive effects of DNA methylation to be somehow retained through this transitory period of low DNA methylation? This merely represents one potential model for how a link between these two modifications may bear out function during this reprogramming environment of low DNA methylation.

### 1.1.3 The biological significance of the RNA methyltransferase enzymes

Limited work has been performed so far on RNMTs in mammalian systems, however from a combination of disease models and mouse knockouts potential roles for these enzymes can be hypothesised. These shall be considered for each member of this family of enzymes in accordance with currently published experimental data.

#### 1.1.3.1 Nucleolar protein 2 (NOP2) is necessary for cellular proliferation

The RNMT NOP2, also known as NOL1, NOP120, NSUN1, and P120, was initially discovered as a proliferation associated nucleolar antigen present in HeLa cells (Freeman et al., 1988). NOP2 overexpression correlates sufficiently strongly with poor cancer prognoses that it was developed as a diagnostic biomarker (Saijo et al., 2001; Bantis et al., 2004). Knockdown of *Nop2* in early mouse embryos results in embryos which are unable to develop beyond the morula stage to form a blastocyst (Cui et al., 2016). NOP2's expression is typically restricted to proliferating cells and is cell cycle dependent, with a large increase in both protein and mRNA levels during late G1 and S phase in human fibroblasts (Fonagy et al., 1993). Knockdown of *Nop2* in mitogen-induced human lymphocytes is sufficient to cause cell cycle arrest, with cells unable to enter S phase (Fonagy et al., 1992). This is most likely due to defective Cyclin D2 expression, as NOP2 can bind to the Cyclin D2 promoter T-cell factor binding element to induce gene expression (Hong et al., 2016). *Nop2* is also found in the locus responsible for Cri du chat syndrome, a severe disorder associated with improper brain development (Wu et al., 2005).

#### 1.1.3.2 NOP2/Sun domain protein 2 (NSUN2) and TRDMT1 are synergistically involved in differentiation

NSUN2 is one of the most upregulated Myc response genes, its expression correlates to cancer aggressiveness, and null mutants have been found to suffer from mental underdevelopment and

microcephaly (Okamoto et al., 2012; Frye et al., 2006; Martinez et al., 2012). *Nsun2* knockout mice show reduced size and exhibit a failure of terminal differentiation in a number of tissues, particularly in the gonads, causing sterility in males (Blanco et al., 2011). *Trdm1* knockout mice display no phenotype, however differentiation defects are observed in zebrafish knockouts in a number of organs including the retina, liver, and brain (Goll et al., 2006, Rai et al., 2007). Double knockouts of *Trdm1* and *Nsun2* do however result in a high proportion of synthetic lethality and more severe phenotypes when compared at E14.5 (Tuorto et al., 2012) (Fig. 1.3A, Fig. 1.3B). These experiments and clinical observations indicate key roles of the developmental and proliferative programs and would be consistent with phenotypes related to differentiation, which is intrinsically linked to DNA methylation. Additionally, they show that there is a capacity for compensation between the RNMTs, which may mask the phenotypes observed in single enzymatic mutants.

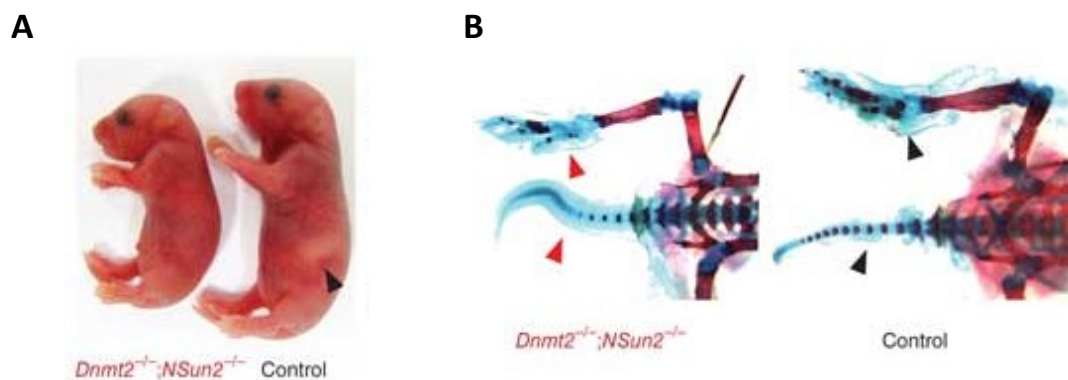


**Fig. 1.3 Morphology of *Trdm1* and *Nsun2* double knockout embryos reveal global developmental defects**

**(A)** Newborn *Dnmt2*<sup>-/-</sup>;*NSun2*<sup>-/-</sup> mice are smaller than control littermates (shown: *Dnmt2*<sup>-/-</sup>;*NSun2*<sup>-/-</sup>), and many do not develop a milk spot (indicated by the arrowhead) and die at P3. Adapted from Tuorto et al., 2012. Figure reproduced with permission of the rights holder, Springer Nature Publishing AG.

**(B)** Ventral view of tail and right leg from littermates. Mineralized bones and cartilage are stained red and blue, respectively. Arrowheads indicate incomplete skeleton ossification in newborn *Dnmt2*<sup>-/-</sup>;*NSun2*<sup>-/-</sup> mice (control littermate shown: *Dnmt2*<sup>-/-</sup>;*NSun2*<sup>-/-</sup>). Adapted from Tuorto et al., 2012. Figure reproduced with permission of the rights holder, Springer Nature Publishing AG.

#### 1.1.3.3 NOP2/Sun domain protein 4 (NSUN4) and NOP2/Sun domain protein 5 (NSUN5) are involved in mitochondrial translation and neurodevelopment respectively.

NSUN4 is located in the mitochondria where it methylates 12S ribosomal RNA at the cytosine 911 position and forms a complex with mitochondrial elongation factor 4 (mtEF4) which is required to assemble the small mitochondrial ribosomal subunit (Metodiev et al., 2014). Global knockout of *Nsun4* causes embryonic lethality, while tissue-specific knockout in the heart results in cardiomyopathy and dysfunctional mitochondria (Metodiev et al., 2014). The electron transport chain complexes with proteins encoded on the mitochondrial DNA showed substantial downregulation at 20 weeks post-birth (complex I, III, IV and V), while the entirely genomic DNA encoded complex II showed increased expression (Metodiev et al., 2014). This is consistent with a global mitochondrial translation defect. *Nsun5* is located in the locus associated with Williams-Beuren syndrome (WBS), a neurodevelopmental disorder characterised by a range of cognitive deficiencies and mild to moderate mental retardation (Li et al., 2009). These phenotypes can be recapitulated in mice upon the knockdown of *Nsun5* (Zhang et al., 2019). While the function of NSUN4 appears restricted to its location within the mitochondria, NSUN5 knockout displays a developmental phenotype, as has been outlined above for NOP2, NSUN2, and TRDMT1.

#### 1.1.3.4 RNA methyltransferases in chromatin reorganisation

Perhaps most surprisingly, last year multiple RNMTs were implicated in causing chromatin reorganisation in leukaemia cell lines. Such a role would enable information transfer from RNA transcripts to genomic loci, which to date has only been robustly characterised for piRNA in germ cells, R-loops, and in the plant RNA-dependent RNA polymerase (Aravin et al. 2008, Skourti-Stathaki et al., 2014, Xie et al., 2004). NOP2 interacts with BRD4 to directly recruit RNA polymerase II, while NOP2/Sun domain protein 3 (NSUN3) and DNMT1 interact with

hnRNPK to recruit the transcription factors GATA binding protein 1 (GATA1) and SPI1/PU.1 (Fig. 1.4A, Fig. 1.4B) (Cheng et al., 2018). These transcription factors in turn recruit positive transcription elongation factor b (pTEFb) which leads to RNA polymerase II assembly and catalyses the phosphorylation of DRB sensitivity inducing factor (DSIF), negative elongation factor (NELF), and RNA polymerase C-terminal domain (CTD) serine-2, which are necessary to form elongating RNA pol II (Fig. 1.4B) (Cheng et al., 2018, Zhou et al., 2012, Yamada et al., 2006). These conclusions were drawn primarily from a number of western blotting, IP, co-IP and colocalisation IF experiments. While this does not provide enough information to confirm the temporal sequence outlined above, it emphasises the capacity for a system with these defined interactions to enable crosstalk between epitranscriptomics and transcription and epigenetics (Cheng et al., 2018).

A common theme in the phenotypic characterisations of RNMT knockouts, and disease mapped to these genomic loci, are proliferative and differentiation defects. This would be consistent with functionalities related to the DNMTs, although not uniquely so. This body of work analyses the role of these RNMTs in relation to DNA methylation and development more broadly. The RNA context in which 5mC modifications occur is critical to understand their biological significance, with the current understanding reviewed below.

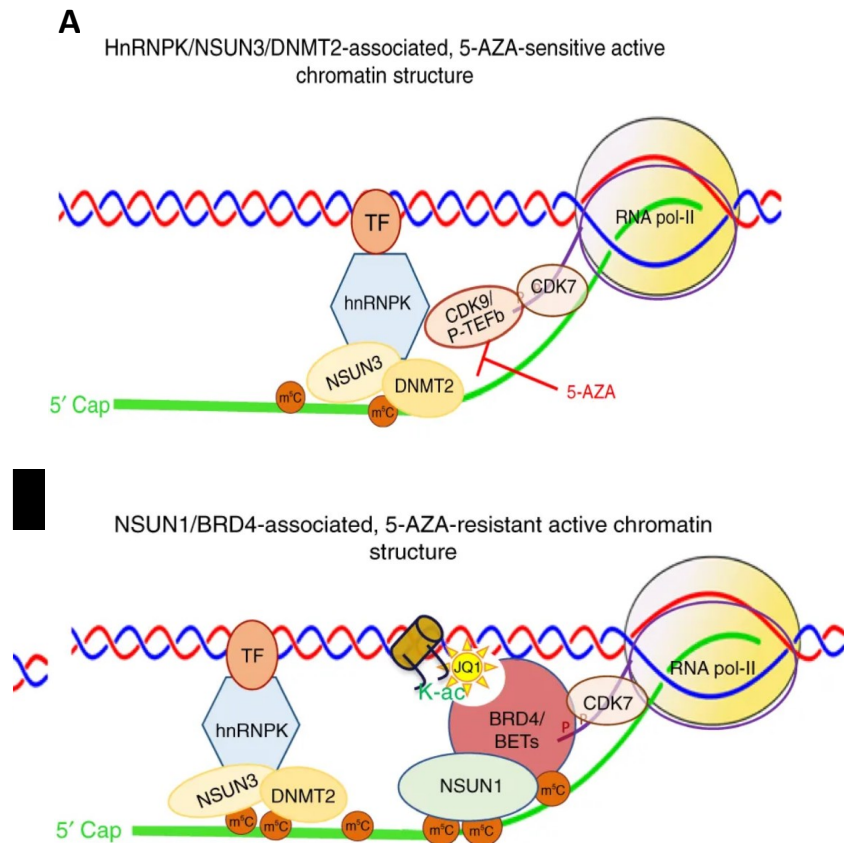


**Fig. 1.4 The reported roles of RNMTs in chromatin reorganisation and transcriptional upregulation.**

Representations of NSUN3/DNMT-associated **(A)** and NOP2-associated **(B)** active chromatin structure. **(A)** NSUN3 and DNMT1 interact with hnRNPK to recruit the transcription factors GATA1 and SPI1/PU.1 **(B)** NOP2 interacts with BRD4 to directly recruit RNA polymerase II. Adapted from Cheng et al., 2018. Figure reproduced under a Creative Commons Attribution (CC-BY) Licence.

### 1.1.4 RNA 5mC characterisation

Having considered the phenotypic outcomes of disrupting RNMT function, it is critical to address the current understanding of where these modifications exist within the transcriptome. This shall be introduced by considering each RNA subtype in turn in which 5mC has been characterised to date.

#### 1.1.4.1 5mC in rRNA maturation and stress response

The functionality of 5mC in RNA is generally unknown, however there are a number of characterised instances in which biological significance can be ascribed. Depletion of NOP2 in yeast has been shown to inhibit 27S pre-rRNA maturation, which may help to explain its expression through the cell cycle, correlation with cell proliferation, and lethality upon knockdown in early mouse embryos (Hong et al., 1997). NSUN5, much like NOP2 and NSUN4, has been found to methylate rRNA (Sharma et al., 2013). Loss of conserved NSUN5-dependent RNA 5mC methylation at C2278 in *S. cerevisiae* 25S rRNA upon *Nsun5* knockout has been shown to alter the conformation adopted in this region of the ribosome, which is in close proximity to the peptidyl transferase site (Schosserer et al., 2015). This results in reduced translational fidelity and a shift towards favoured recruitment of oxidative stress-responsive mRNAs to the polysome resulting in a phenotypic increase in stress resistant and organismal life expectancy in *C. elegans*, *S. cerevisiae*, and *D. melanogaster* upon *Nsun5* knockdown (Schosserer et al., 2015). Extended lifespan can be recapitulated by mutation of C2278 to G2278, which cannot be methylated by NSUN5, indicating that the presence or absence of the 5mC modification is the mechanistic determinant of the phenotype (Schosserer et al., 2015).

#### 1.1.4.2 5mC in tRNA stability, genetic recoding, and stress response

Much of the functional characterisation of RNA 5mC has focused on tRNAs, which is in no small part due to the high abundance of these transcripts, coupled with historical technical limitations. More recently crosslinking experiments involving NSUN2 and TRDMT1 showed a large proportion of the detected catalytic activity to be directed to tRNA species in HEK293 and HeLa cell lines respectively (Hussain et al., 2013; Khoddami, et al., 2013). The mature form of a leucine-inserting amber suppressor tRNA (SUP53) in *S. cerevisiae* is known to have a 5mC modification at the wobble position (5mC34). Genetic manipulation to generate a hypomethylated tRNA at the wobble position correlates with reduced repressor function, indicating a possible role for 5mC in codon recognition and translational efficiency (Strobel et

al., 1986). Developing from this, it has been recently postulated that RNA 5mC could form the basis for genetic recoding, as has been observed in *E.coli* for whom an insertion of 5mC into the 2nd position of the CCC codon led to 4% recoding of proline to leucine (Hoernes et al., 2016).

Post-transcriptional cytosine methylation functionalities in tRNA are not necessarily restricted to bases within the anticodon. Chemically synthesised anticodon stem and loop regions of *S. cerevisiae* phenylalanine tRNA with or without 5mC on position C40 have been shown to adopt different conformations in a Mg<sup>2+</sup>-dependent manner by both circular dichroism and NMR (Chen et al., 1993). This indicates that modifications located outside of the anticodon loop have the potential to modulate functionality. Such effects have been observed for other nucleoside modifications, such as 7mG. Reversible hydride transfer to 7mG47 results in a less ordered tRNA structure, while *E.coli* tRNA<sup>Phe</sup> and tRNA<sup>Ala</sup> display a 10-fold variation in V<sub>max</sub> for the aminoacyl reaction with/without *in vitro* methylation to 7mG10 (Arcari et al., 1978; Roe et al., 1973; Bjork et al., 1987).

*Trdmt1* and *Nsun2* single and double knockout mice display reduced steady state levels of their specific target tRNAs, which correlates with a reduction of the rate of protein synthesis. This shows 5mC to be critical for the stability of certain tRNAs (Tuorto et al., 2012). Further, 5mC-dependent tRNA cleavage has been described as a mechanism for global translational control in stress conditions. NSUN2 is inhibited by UV irradiation in human and mouse keratinocytes. Upon UV exposure, or direct *Nsun2* knockout, tRNAs are cleaved to produce 5' and 3' products leading to reduced translation rates and activation of stress pathways. The 5' fragments generated can inhibit translation if transfected or injected into unstressed cells, indicating that they have direct translational inhibition properties separate from a simple reduction of the steady-state level of full-length tRNAs (Yamasaki et al., 2009; Blanco et al., 2014). Similar behaviour has been detailed for TRDMT1, with overexpression of TRDMT1 in mouse embryonic fibroblasts

(MEFs) sufficient to protect against tRNA cleavage in stress conditions (Schaefer et al., 2010). Ribosome profiling has shown that *Nsun2* knockdown, as well as subsequent rescue, in hair follicle stem cells (HFSCs) alter ribosome occupancy along mRNA transcripts. Differing changes were observed when genes in separate gene ontology categories were assessed. NSUN2 knockdown appears to act to upregulate translation of stress response, DNA damage response, and cell migration genes, while downregulating genes for differentiation, endocytosis, and keratinization (Blanco et al., 2016). This highlights nuances in the global 5mC-dependent translation control mechanism, whereby translational efficiency is broadly reduced by stress and/or NSUN2 downregulation, however the translation of genes associated with the stress response are increased.

#### 1.1.4.3 5mC in mRNA regulates transcript degradation, translational efficiency, and nuclear export.

Although tRNAs represent the most enriched RNA subtype for 5mC, ~83% of all candidate 5mC sites identified in a comprehensive bisulphite screen of the HeLa transcriptome were located within mRNAs (Squires et al., 2012). mRNA therefore represents the bulk of 5mC candidate modification sites, despite being proportionally devoid of this modification. Within mRNAs there is an enrichment of 5mC in both the 5' and 3' untranslated regions (UTRs) compared to values predicted from a random distribution. A substantial association between 5mC candidate sites in mRNA and Argonaute I-IV protein binding sites was observed (Squires et al., 2012). This link to miRNA regulation through enrichment in Argonaute binding sites was found by correlating the bisulphite data to photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation sequencing (PAR-CLIP-seq) data (Squires et al., 2012, Hafner et al., 2010). Separate later work using bisulphite sequencing of poly(A) RNA from mouse ES cells (mESCs) and brain tissue has shown a similar pattern of enrichment of 5mC in mRNA 5' UTRs and 3' UTRs, but additionally characterised an enrichment in 5mC around the

translation start site (Amort et al. 2017). However, this dataset indicates a depletion of 5mC in miRNA target sites, contrasting with the work of Squires et al. (Amort et al. 2017). Computational prediction of RNA binding protein sites was used by these investigators, and so a disconnect between experimentally-verified crosslinking analysis and these computational predictions may explain the discrepancies (Murakami et al., 2010). Together, these results indicate that 5mC is very widespread in mRNA, particularly in the 5' and 3' UTRS, and may have regulatory roles, potentially associated with the miRNA pathway. More recently higher stringency workflows for identifying mRNA 5mC sites in human and mouse tissue have been developed (Huang et al., 2019). This methodology substantially reduces the number of mRNA 5mC sites detected. However, as with all analyses of this nature, there is an open philosophical question over the balance between stringency and false negative rate. The sites that are excluded by this more stringent protocol are enriched for being clustered together, which could be argued to imply that they were initially called due to incomplete conversion during bisulphite treatment (Huang et al., 2019, Amort et al. 2017, Squires et al., 2012). Equally clusters of lowly methylated sites could be bona fide targets which require a lower frequency of methylation on any given site relative to isolated 5mC sites to give functionality. Further study is necessary to understand the contextual relevance of 5mC modifications in mRNA, which could then be fed back to guide stringency cut-offs. These developments only further highlight how techniques and analysis pipelines continue to open up novel avenues through which to study RNA modifications.

By interactome capture NOP2, NSUN2, NSUN4, and NSUN5 all bind to mRNA, and so represent putative RNMTs to mRNA (Castello et al., 2012). Indeed, knockdown of NSUN2 reduces mRNA 5mC methylation with NSUN2-dependent mRNA 5mC sites sharing sequence and predicted structural motif information with NSUN2-dependent tRNA modifications (Huang et al., 2019). Interestingly, in HeLa cells knockdown of NSUN2, but not NOP2, NSUN5 or NOP2/Sun domain protein 6 (NSUN6) resulted in lower 5mC levels in mRNA, but not total

RNA, as quantified by LC-MS/MS (Yang et al., 2017). Both *in vitro* and *in vivo* experiments have found that 5mC in the coding sequence of mRNAs results in reduced translational efficiency by western blotting and ribosome profiling respectively (Hoernes et al., 2016, Huang et al., 2019). This has been observed both for the addition of 5mC to novel mRNA sites *in vitro* and the removal of biological mRNA 5mC sites through *Nsun2* knockout (Hoernes et al., 2016, Huang et al., 2019). When transcript-specific analyses have been performed for mRNAs, modulation of transcript decay and translation efficiency of p16INK4 and *Cdk1* mRNA respectively have been observed (Zhang et al., 2012; Xing et al., 2015). Interestingly, NSUN2 binding appears to increase the half-life of p16INK4 mRNA without changing translation efficiency, while NSUN2-catalysed methylation of CDK1 mRNA at C1733 increases translation efficiency without changing mRNA half-life. Methylated-individual-nucleotide-resolution crosslinking and immunoprecipitation (miCLIP) does not support a direct role for NSUN2-mediated 5mC in regulation of the steady-state levels of its targets, however it would be interesting to apply this methodology to some of the other RNMTs and their respective targets (Hussain et al., 2013).

A specific mRNA 5mC reader protein has recently been discovered and characterised. Aly/REF export factor (ALYREF) is a nuclear export protein which preferentially binds to 5mC-containing mRNAs (Yang et al., 2017). In shuttling between nucleus and cytoplasm it facilitates targeted mRNAs export in a NSUN2-dependent manner (Yang et al., 2017). Rescue of ALYREF and NSUN2-target mRNA retention in the nucleus upon *Nsun2* knockout can be induced by expression of wild type, but not catalytically dead, NSUN2 (Yang et al., 2017). Non-methylated RNAs were not altered in their localisation by either knockdown of *Nsun2* or *Alyref*, while NSUN2-independent 5mC containing mRNAs still showed increased nuclear retention upon *Alyref* knockdown (Yang et al., 2017). Together, this strongly implicates 5mC mRNA methylation in the proper regulation of ALYREF-dependent mRNA export from the nucleus (Yang et al., 2017).

RNA 5mC modifications have been functionally characterised in rRNAs, tRNAs, and mRNAs and ascribed to a number of functionalities including translational regulation as well as transcript biogenesis and decay. The putative functions of RNA 5mC are summarised in Table 1.1.

### 1.1.5 RNA 5hmC characterisation

#### 1.1.5.1 RNA 5hmC is intrinsically linked to 5mC

5mC functionalities may be dynamically controlled via erasure as well as deposition. Such a mechanism may not be as necessary in RNA as it is in DNA, as RNA degradation could theoretically bring about an analogous effect. Although insights into the presence or nature of such a hypothesised erasure are lacking, hypothesis can be driven indirectly from the emerging, complex picture of the mechanism(s) by which 5mdC is removed during epigenetic reprogramming, which is/are in themselves poorly understood (Hill et al., 2014). The ten eleven translocation (TET) family of dioxygenases can catalyse the sequential oxidation of 5mdC to hydroxymethyl- (5hmdC), formyl- (5fdC) and finally carboxyldeoxycytidine (5cadC) (Tahiliani et al., 2009; Ito et al., 2011; He et al., 2011). It has been shown that 5cadC may be removed by thymine-DNA glycosylase (TDG) as part of the base excision DNA repair (BER) pathway, and thus this has been proposed as a possible mechanism of DNA demethylation (Hajkova et al., 2010, He et al., 2011, Santos et al., 2013). The expression of both TET1 and TET2 increases during epigenetic reprogramming, which lends credence to such a theory (Koh et al., 2011). However, in both zygotic and primordial germ cell epigenetic reprogramming, the kinetics of the appearance of 5hmdC do not facilitate such a mechanism as the primary route for the loss of 5mdC, (Hill et al., 2014, Amouroux et al., 2016, Hill et al., 2018). Rather, it appears that 5hmC and its derivatives act to prevent the immediate reestablishment of DNA methylation, and thus are required to maintain, but not to generate, DNA hypomethylation (Amouroux et al., 2016;

Hill et al., 2018). Interestingly, 5hmC, 5fC, and 5caC has also been observed in RNA and found to be generated by the same TET enzymes (Fig. 1.5) (Fu et al., 2014). 5hmC is intrinsically linked to its obligatory precursor 5mC, and thus it is important to investigate the biochemistry of both to fully understand the role(s) of either.

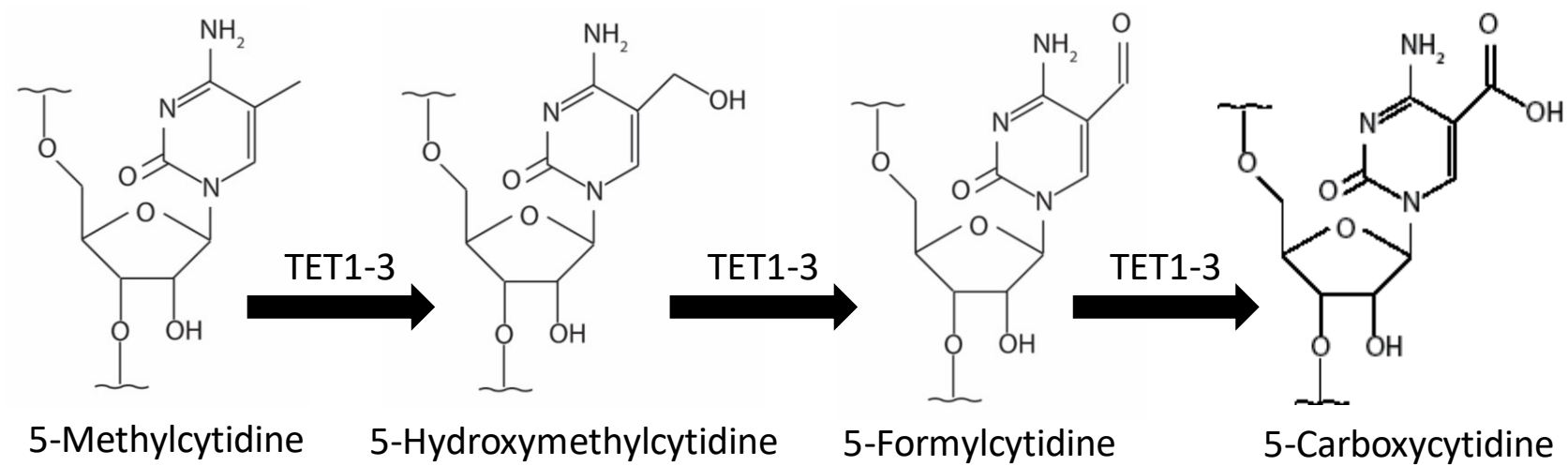


**Fig. 1.5** TET-mediated oxidations of RNA 5mC.

The progressive oxidation of RNA 5-methylcytosine to 5-carboxycytidine is displayed in skeletal formulae. Wiggly lines indicate the extended ribonucleotide backbone.

#### 1.1.5.2 RNA 5hmC increases translational efficiency and is necessary for development

It has been shown that RNA 5hmC is present in all 3 domains of life and is enriched 40-fold in polyA-containing RNA in HEK193T cells (Huber et al., 2015). Considering the relative depletion of the amount of 5mC in mRNA species, such a strong enrichment of 5hmC in poly(A) RNA may indicate that 5mC within mRNAs is more dynamic than those in tRNA or rRNA. This would be an interesting scenario if verified, as dynamic deposition and erasure would strongly implicate regulatory function(s). Within the brain tissue of mice 5hmC is present at ~0.001% of total cytosine nucleosides in RNA, representing a 1000x less abundant modification than 5mC (Huber et al., 2015). The abundance of 5hmC in RNA correlates with RNAs displaying high ribosome occupancy in *Drosophila* S2 cells (Delatte et al., 2016). Furthermore, *in vitro* translation of *in vitro* transcribed RNAs containing unmodified, methylated, or hydroxylated luciferase mRNA show that, while RNA 5mC reduces protein yield, 5hmC can restore translation levels to that of the unmodified transcript (Delatte et al., 2016). Accordingly, 5hmC may function through the removal, or masking, of 5mC in addition to any functions actively brought about by 5hmC. The phenotypes displayed by *Tet* knockouts will be summation of the effects of the loss of DNA 5hmdC and RNA 5hmC. To date, the closest experiments to disentangling these processes in an organism-wide context is work performed in *drosophila*. *Drosophila* display very low levels of DNA methylation and so knockout of the single *Tet* locus in their genome has been hypothesised to mostly result in RNA 5hmC defects over 5hmdC effects. This thought process does involve a rather critical assumption. This results from *drosophila* containing non-zero 5mdC DNA methylation (~0.1%–0.6% of all deoxycytidine), which has been found to have functionality in maintaining retrotransposon silencing (Lyko et al., 2000, Gowher et al., 2000, Phalke et al., 2009). With this caveat in mind, when the *Tet* locus is deleted *drosophila* embryos are no longer viable and are unable to progress beyond the late pupal stage. Tissue specific knockouts in neurones, heart muscle, and ovaries showed TET to

be required for proper development and function of each of these structures (Wang et al., 2018). This indicates, but not conclusively so, that RNA 5hmC is necessary for multiple developmental programs. The putative functions of RNA 5hmC are summarised in Table 1.1.

Although by no means comprehensive, what information is known about the role(s) of RNA 5mC and 5hmC give interesting tentative signs of potential key regulators of gene expression circuitry. How these emerging functions, among other possible roles, may be perturbed during epigenetic reprogramming could elucidate further their significance *in vivo*.

### 1.1.6 RNA m6A writers, readers, and erasers: The molecular machinery used to regulate differentiation

6-methyladenine (m6A) is an enzymatically generated RNA modification whose biological function is substantially better understood than 5-methylcytidine. In mammalian cells it is present at approximately 3.0 sites per mRNA transcript (Perry et al., 1975). Reader, writer, and eraser proteins for m6A have been characterised, providing the full repertoire of machinery for a reversible modification whose presence can confer function (Fig 1.6) (Liu et al., 2014).

#### 1.1.6.1 Writers of m6A

The methyltransferase-like (METTL) family of enzymes specifically methylate the nitrogen at the 6' position of the adenine heterocyclic purine ring. The first members of this family to be found to generate this modification *in vivo* are METTL3 and METTL14, which form a complex with the adapter proteins and splicing factors WTAP and KIAA1429 (Wang et al., 2014, Liu et al., 2014, Schwartz et al., 2014). The two methyltransferases display synergistic effects on enzymatic activity, with crystal structures showing METTL14 acting as an RNA-binding platform with METTL3 acting as the primary catalytic core (Wang et al., 2016, Liu et al 2014). The METTL3-METTL14 complex has been shown to act co-transcriptionally on thousands of

transcripts in mouse ES cells (Knuckles et al., 2017). Targeting is primarily to protein coding transcripts, although non-coding transcripts including snoRNA and miRNA clusters were also identified. Intriguingly, modulation of miRNA seed sequence has been shown to be sufficient to direct ab initio m6A deposition by METTL3 in multiple mouse stem cell lines, indicating a miRNA-directed targeting mode for m6A methylation (Chen et al., 2015).

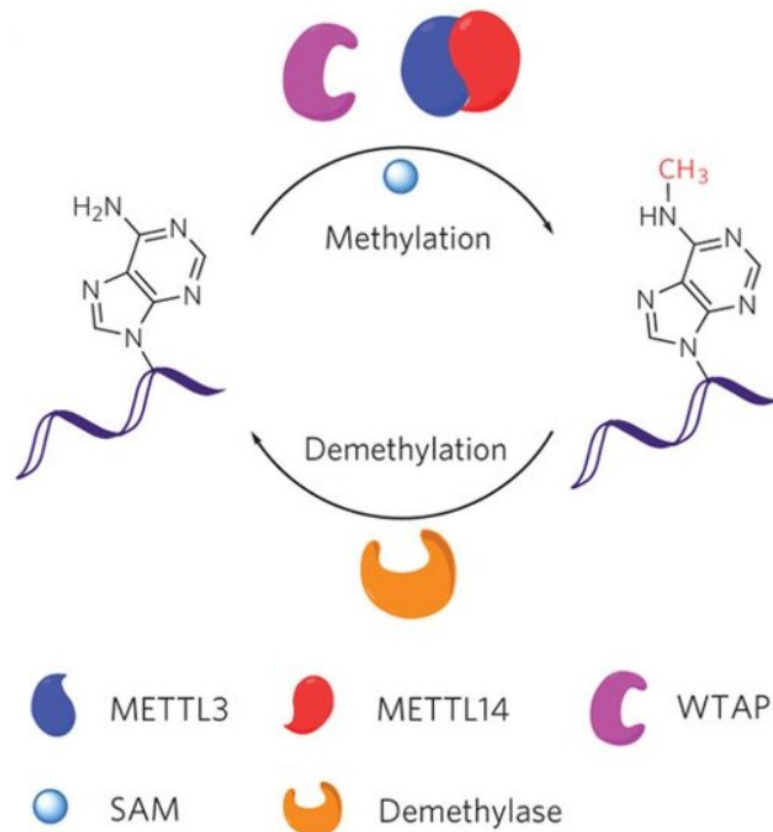


**Fig 1.6 Reversible m6A methylation.**

The reversible formation of 6-methyladenosine is displayed. The METTL3-METTL14-WTAP complex is shown as a major m6A reader complex, along with the SAM cofactor required for methylation. A demethylase enzyme is displayed catalysing the reverse reaction. Adapted from Liu et al., 2014. Figure reproduced with permission of the rights holder, Springer Nature Publishing AG.

Knockdown and sequencing experiments found that WTAP-dependent and WTAP-independent m6A sites exist in *p53* knockout mouse embryonic fibroblasts (Schwartz et al., 2014). This indicated that other m6A methyltransferases beyond METTL3 and METTL14 were acting in these cells. METTL16 has more recently been discovered as a m6A

methyltransferase which is key to S-Adenosylmethionine (SAM) homeostasis (Pendleton et al., 2017). The mRNA of the SAM synthases gene *Mat2a* is methylated by METTL16 on intronic hairpins in high SAM conditions to prevent U6 RNA recruitment leading to intron retention and transcript degradation (Pendleton et al. 2017, Shima et al., 2017). The crystal structure of the N-terminal region of METTL16 has been determined, revealing a positive-charged groove used to bind substrate RNA opening towards a disordered region around the catalytic NPPF motif (Mendel et al., 2018). Using the *Mat2a* hairpin structure as a bona fide *in vivo* METTL16 target sequence-structure context, multiple *in vitro* RNA transcripts were designed to probe the critical aspect for METTL16-dependent methylation (Mendel et al., 2018). A stem region of at least 3 base pairs is necessary for activity, which can be rescued by an artificial C/G stem, indicating that RNA secondary structure is necessary for target recognition in addition to the primary nonamer sequence (UACAGAGAA) around the methylation site (Mendel et al., 2018). With this information, it would be interesting to use *in silico* approaches to attempt to identify METTL16 target sites throughout the transcriptome. Comparing the wild type, heterozygous and homozygous *Mettl16* knockout transcriptomes of E2.5 and E3.5 mice reveals only 4 genes with altered transcript abundance at E2.5, which increases to several thousand one day later at E3.5 (Mendel et al., 2018). One of the 4 genes is the previously identified target *Mat2a* gene encoding SAM synthase (Mendel et al., 2018). It is likely that the transcript-wide defects and following lethality are therefore directed by secondary effects of *Mettl16* knockdown, namely a lower SAM pool leading to methylation defects throughout all cellular processes (including protein, DNA, RNA, and metabolites). Clearly at least two distinct m6A deposition machineries exist in mammals targeting separate sites within the transcriptome.

#### 1.1.6.2 Functions of m6A independent of reader proteins

N6-methyladenosine is enriched within the 5' and 3' UTR regions, 3'-most exons and long exons in mRNA (Meyer et al., 2012, Dominissini et al., 2012, Batista et al., 2014, Schwartz et al., 2014,

Ke et al., 2015). It has been shown that there is an increase in m6A around the stop codons, opening up the possibility of a role in translational termination (Meyer et al., 2012, Dominissini et al., 2012, Batista et al., 2014, Schwartz et al., 2014). However, more recent conflicting results indicate that this may not necessarily be the case (Ke et al., 2015).

Small alkyl groups such as methyl groups are not very reactive. However, through steric hindrance m6A can disrupt base-pairing and so alter RNA secondary structure. Such thermodynamic destabilisation has been observed by UV melting for *in vitro* synthesised m6A modified RNAs (Kierzek et al., 2003). This has been shown to be the basis of a riboswitch in the lncRNA metastasis-associated lung adenocarcinoma transcript (MALAT1) in HeLa cells (Liu et al., 2015). m6A-dependent disruption of a hairpin with an opposing U-tract is necessary for heterogeneous nuclear ribonucleoprotein C (HNRNPC) binding (Liu et al., 2015, Ehresmann et al., 1987). This indicates that functionality of this modification does not always require direct m6A-reader proteins, but rather effects can be brought about through secondary structure modulation.

Additionally, the characterised METTL enzymes do not exist in isolation *in vivo*. As has been outlined above, METTL3 and METTL14 exist in a complex with WTAP and KIAA1429, both of which acts as splicing factors (Wang et al., 2014, Liu et al., 2014, Schwartz et al., 2014). Given that this complex has been shown to be co-transcriptionally recruited to pre-mRNA transcripts, it is consistent that m6A methylation may be involved in regulating splicing transcriptome-wide (Knuckles et al., 2017). METTL16 contains two vertebrate conserved regions (VCRs) which are sufficient to direct splicing of the *Mat2a* pre-mRNA transcript, while HA-tagged METTL16 interaction partners, as identified by mass spectrometry, are enriched for splicing factors (Pendleton et al., 2017, Mendel et al., 2018). While METTL16 may also modulate splicing, there are no global splicing defects upon homozygous *Mettl16* knockout in E3.5 mouse embryos (Mendel et al., 2018).

### 1.1.6.2 Readers of m6A

YTHDF1, YTHDF2, and YTHDC1 have all been identified as proteins capable of m6A-dependent RNA binding. YTHDF1 increases translational efficiency by recruiting mRNA containing m6A to the ribosome and by increasing the rate of initiation of translation (Wang et al., 2015). YTHDF2 has been shown to bind over 3000 transcripts in HeLa cells with a sequence preference for G(m6A)C (Wang et al., 2014). It binds m6A-containing mRNAs with its C-terminal domain, while the N-terminal domain acts to recruit RNA to processing bodies for degradation by decapping and subsequent 5'-3' exoribonuclease activity (Wang et al., 2014, Sheth et al., 2003). YTHDF2 also recruits the carbon catabolite repressor 4 – negative on TATA (CCR4-NOT) deadenylase complex which removes the polyA tail of mRNAs leaving them exposed to 3'-5' exoribonucleases (Du et al., 2016). At first glance, YTHDF1 and YTHDF2 appear to have functionally antagonistic effects, resulting in either increased translation efficiency or decreased half-lives of m6A containing mRNAs, particularly as they show ~50% overlap in their target transcripts (Wang et al., 2015). However, what this system represents is a collection of machineries which enables transient sharp spikes in protein output. This would be of benefit for rapid cell state transitions (Fig. 1.7). The current thinking in the field has coalesced around this as a unifying theory.

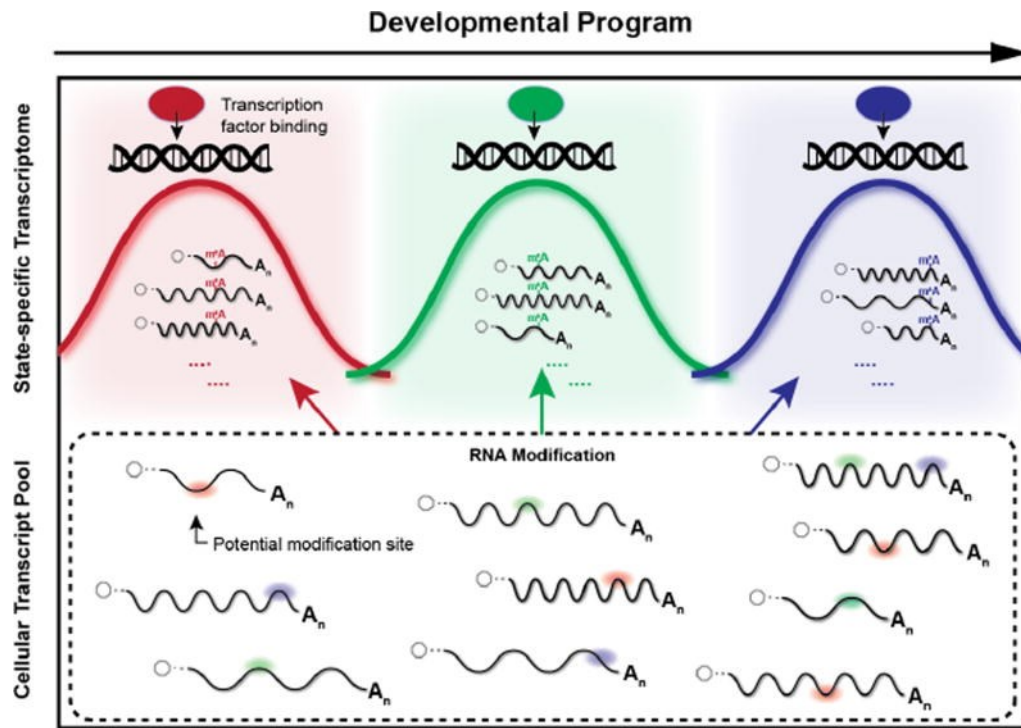


**Fig 1.7 RNA m<sup>6</sup>A modifications groups transcripts for cell fate transitions.**

Developmental programs require rapid switching of the cellular transcriptome to bring about phenotypic changes. Recruitment of transcription factors can alter the composition of the RNA pool by introducing required transcripts to the existent mRNA population. In order to accomplish more rapid transitions, cells utilize the existing pool of mRNA, tuning their expression accordingly. One way to differentiate distinct groups of mRNAs from a diverse cellular pool is by post-transcriptional modification by m<sup>6</sup>A, which marks transcripts for incorporation into pathways for translation and decay (Roundtree et al., 2017). Figure adapted with permission of the rights holder, Elsevier.

YTH domain-containing protein 1 (YTHDC1) is a m<sup>6</sup>A reader protein which has been implicated in METTL16-catalysed methylation sites, as well as those generated by METTL3. YTHDC1 is responsible for the degradation of *Mat2a* mRNA with m<sup>6</sup>A modifications on the intronic region retained in high SAM conditions (Shima et al., 2017). As the *Mat2a* transcript encodes the SAM synthase gene, this provides negative feedback to regulate the levels of SAM synthase in the cell. X-inactive specific transcript (XIST) lncRNA has at least 78 m<sup>6</sup>A modification sites which are necessary for the repression of the inactive X chromosome (Patil et al., 2016). In accordance with this knockdown of METTL3 resulted in derepression of genes such as *Gpc4* and *Atrx*. This is brought about by YTHDC1 binding to these sites, with tethering of YTHDC1 to XIST sufficient to rescue X inactivation upon loss of m<sup>6</sup>A by METTL3

knockdown (Patil et al., 2016). These experiments show YTHDC1 recruitment to be necessary and sufficient for X chromosome inactivation. It follows that YTHDC1, targeted to XIST RNA through the abundant m6A modifications on this transcript, facilitates transcription repression. This may be through recruiting transcriptional silencers, which would be in keeping with miCLIP experiments showing SHARP, LBR, HNRNPU and HNRNPK to all be YTHDC1 interaction partners (Xiao et al. 2016).

#### 1.1.6.3 Erasers of m6A

N<sup>6</sup>-methyladenosine can be enzymatically removed from transcripts, giving rise to a truly reversible modification. Both Fat mass and obesity-associated protein (FTO) and AlkB homolog 5 (ALKBH5) have been characterised as RNA m6A demethylases *in vitro* and *in vivo* (Jia et al., 2011, Zheng et al., 2013, Choi et al., 2019). *Fto* knockout mice only display altered m6A peaks in a subset of peaks by methylation RNA immunoprecipitation sequencing (MeRIP-Seq), indicating that the majority of m6A sites in the transcriptome are not targeted by this demethylase (Hess et al., 2013). This could mechanistically explain why SNPs in the *Fto* locus 16q12.2 correlate with increased body mass index (BMI) and obesity risk in multiple human populations as opposed to phenotypes related to differentiation, development, and germline formation, as is the case for the majority of m6A readers, writers and erasers (Saunders et al., 2007, Fawcett et al., 2010). These SNPs have attracted research interest, as they map to one of the highest ranking bins for obesity and BMI in a comprehensive human genome-wide association study (GWAS) involving over 31,000 individuals (Saunders et al., 2007, Fawcett et al., 2010). Mutations in the gene are strongly correlated with increased energy intake, but not to any variations in expenditure, leading to the prevailing view that feeding behaviour is a key element to FTO deficiency-induced adiposity (Haupt et al., 2009). In rats upregulation of FTO in the hypothalamus, which is

responsible for feeding behaviour, reduced food intake, while knockdown increased food intake (Tung et al., 2010). However, mouse work has shown *Fto* mutants to display reduced adipose tissue, which argues against this model in this system (Fischer et al., 2009). In spite of this, fasting in mice has been shown to decrease FTO expression, which appears to be consistent with the rat and human mutants (Stratigopoulos et al., 2008, Fawcett et al., 2010). The link between FTO and whole-body metabolism has come under scrutiny since the finding that the intronic mutations most commonly found by genome-wide association studies lie within a distal enhancer modulating Iroquois 3 (*Irx3*) gene expression (Smemo et al., 2014). As such, the phenotypes observed due to these mutations are most likely due to increased expression of IRX-3, as IRX-3 deficient mice show reduced body weight (Smemo et al., 2014). While this may provide an alternative explanation for the metabolic phenotypes upon FTO mutation, it does not explain what is observed upon FTO overexpression or modulation in the expression of the gene upon fasting (Tung et al., 2010, Stratigopoulos et al., 2008, Fawcett et al., 2010). Further work is required to further elucidate this branch of m6A metabolism, however FTO-responsive sites are the minority within the transcriptome, and so this represents the exception as opposed to the rule.

ALKBH5 is located primarily in the nucleus, indicating that its m6A demethylating role is most likely direct to maturing and nuclear-retained transcripts (Zheng et al., 2013). Knockdown of *Alkbh5* increases the rate of mRNA transport to the cytoplasm, strongly implicating it in the regulation of mRNA nuclear export (Zheng et al., 2013). It has been identified in demethylating the forkhead box M1 (*Foxm1*) mRNA in the presence of antisense RNA to the *Foxm1* transcript (Zhang et al., 2017). Conceptually, RNA-RNA hybrids may form between the sense and antisense transcripts, potentially revealing that ALKBH5 demethylation requires hitherto unidentified structural elements. Although exciting, further work is required to fully understand the targeting and regulation of ALKBH5 in RNA m6A demethylation. This is particularly

pertinent, as *ALKBH5* knockdown has been shown to slightly reduce the number of m6A sites in glioblastoma stem cells from 39,288 and 34,415, which contradicts the expected results for the knockdown of a demethylase (Zhang et al., 2017). This should be caveated that the m6A RIP-seq method used in this paper is relatively noisy. A less controversial conclusion would be that in this experiment *ALKBH5* knockdown does not result in a clear increase in m6A sites, which, regardless, leads to an incomplete understanding of *ALKBH* function which necessitates additional experimentation.

### 1.1.7 RNA m6A is required for early embryo development

*METTL3* is critical for development, as knockout of *Mettl3* causes early embryo lethality at the post-implantation stage (Geula et al., 2015). Embryos were found to be morphologically normal at E3.5, but by E8.5 embryos had been absorbed (Geula et al., 2015). At E6.5 early differentiation markers had not been activated in the mutants, while knockout of *Mettl3* in mESCs and knockdown in human embryonic stem cells (hESCs) resulted in cells that are unable, or inhibited in differentiation respectively (Batista et al., 2014, Geula et al., 2015). *Mettl3* knockout stem cells also result in the formation of larger, poorly differentiated teratomas *in vivo* relative to wild type cells (Geula et al., 2015). Separate work has found knockdown of *Mettl3* or *Mettl14* in mESCs to have the opposite effect, leading to the loss of self-renewal (Wang et al., 2014). Heterozygous *Mettl16* knockout mice are viable and fertile (Mendel et al., 2018). Homozygous mutants appear at Mendelian frequencies at E2.5 and E3.5, but are completely absent by E8.5, indicating embryonic lethality upon complete gene loss (Mendel et al., 2018). Furthermore, in the unicellular *S.cerevisiae* loss of m6A results in sporulation defects indicating a role in gametogenesis (Clancy et al., 2002).

In mice *YTHDF2* is required for female, but not male fertility, governing oocyte quality (Ivanona et al., 2017). Maternally inherited *YTHDF2* is necessary to clear maternal transcripts and

facilitate the transition from maternal to zygotic gene expression in the early zygote (Ivanona et al., 2017). As a result, *Ythdf2* knockout females give rise to zygotes that do not progress beyond the 2 cell stage (Fig. 1.8) (Ivanona et al., 2017). YTHDC1 is necessary for embryo, sperm and oocyte development (Kasowitz et al., 2018). Global knockouts in mice are unable to reach E11.5 while conditional knockouts in oocytes result in major alternative splicing and alternative polyadenylation defects (Kasowitz et al., 2018).

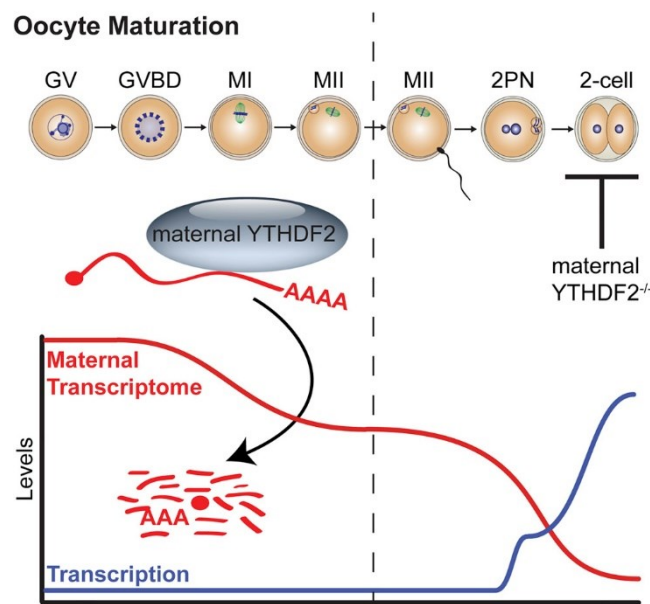


**Fig 1.8 YTHDF2 is required to clear maternal mRNAs and permit zygotic development beyond the 2-cell stage.**

Adapted from Ivanova et al., 2017. Figure reproduced with permission of the rights holder, Elsevier. Figure reproduced under a Creative Commons Attribution (CC-BY) Licence.

Although ubiquitously expressed, ALKBH5 is particularly highly expressed in the testes of mice (Zheng et al., 2013). *Alkbh5* knockout mice are viable, but show defects in spermatogenesis, with reduced sperm number and mobility leading to low fertility rates (Zheng et al., 2013). Knockdown of ALKBH5 in glioblastoma stem cells reduces the self-renewing potential of the cells, as determined by tumoursphere formation frequency in a limiting dilution assay and reduced expression of octamer-binding transcription factor 4 (OCT4), Nanog, and SOX2 (Zhang et al., 2017). In line with this, separate work has shown that knockdown of ALKBH5 enhances differentiation in muscle and adipose models (Choi et al., 2019). Additionally, *Alkbh5* knockdown reduced the proliferation rate of the glioblastoma stem cells (Zhang et al., 2017).

From the genetic manipulation performed to date it is clear that a number of proteins involved in m6A metabolism are critical to viability. Further, phenotypes observed prior to lethality are intrinsically linked to development, gametogenesis, self-renewal and differentiation. 6-Methyladenosine is found in 11 of 17 core pluripotency genes, in addition to many other genes and ncRNAs implicated in pluripotency in mESCs (Fig 1.9) (Batista et al., 2014, Young, 2011). The half-life of m6A modified transcripts in mESCs is on average 2.5 hours shorter than for non-modified transcripts, while the upregulation of expression upon *Mettl3* knockout is positively correlated with the number of m6A sites in wild type mESCs (Batista et al., 2014, Geula et al., 2015). From this it can be hypothesised that the mechanism by which m6A metabolism proteins alter differentiation is by increasing the RNA decay rate of key pluripotency genes. This is consistent with increased global levels of m6A in muscle and adipose differentiation models (Choi et al., 2019). It may well be the case that m6A acts to destabilise the transcriptional networks of a particular cell state to make any given transition more likely, as increasing or decreasing m6A has been shown to improve or reduce the efficiency of *in vitro* dedifferentiation of MEFs respectively (Chen et al., 2015). The role of RNA m6A in development is better characterised than RNA 5mC. However, questions remain, such as how these modifications

may act during epigenetic reprogramming. The functions of RNA m6A are summarised in Table 1.1.

| Modification | Function                                       | RNA target    | Reference                                                                         |
|--------------|------------------------------------------------|---------------|-----------------------------------------------------------------------------------|
| <b>5mC</b>   | Increased / decreased translational efficiency | mRNA          | Xing et al., 2015, Strobel et al., 1986, Hoernes et al., 2016, Huang et al., 2019 |
|              | Stress response                                | tRNA and rRNA | Schosserer et al., 2015, Yamasaki et al., 2009; Blanco et al., 2014               |
|              | Genetic recoding                               | tRNA          | Hoernes et al., 2016                                                              |
|              | Biosynthesis                                   | rRNA          | Hong et al., 1997                                                                 |
|              | miRNA regulation                               | mRNA          | Squires et al., 2012                                                              |
|              | Increased / decreased transcript decay         | mRNA and tRNA | Tuorto et al., 2012, Zhang et al., 2012                                           |
|              | Nuclear export                                 | mRNA          | Yang et al., 2017                                                                 |
| <b>5hmC</b>  | Increased translational efficiency             | mRNA          | Delatte et al., 2016                                                              |
| <b>m6A</b>   | Increased translational efficiency             | mRNA          | Wang et al., 2015                                                                 |
|              | Increased transcript decay                     | mRNA          | Wang et al., 2014, Sheth et al., 2003                                             |
|              | Alternative splicing and SAM homeostasis       | mRNA          | Mendel et al., 2018, Shima et al., 2017                                           |
|              | X-chromosome inactivation                      | lncRNA (XIST) | Patil et al., 2016                                                                |

**Table 1.1 Putative functions of RNA modifications**

Table summarises the putative functions of RNA 5mC, 5hmC, and m6A as outlined in this chapter. The distinct target RNA subtypes are shown in addition to key references reporting these functionalities.

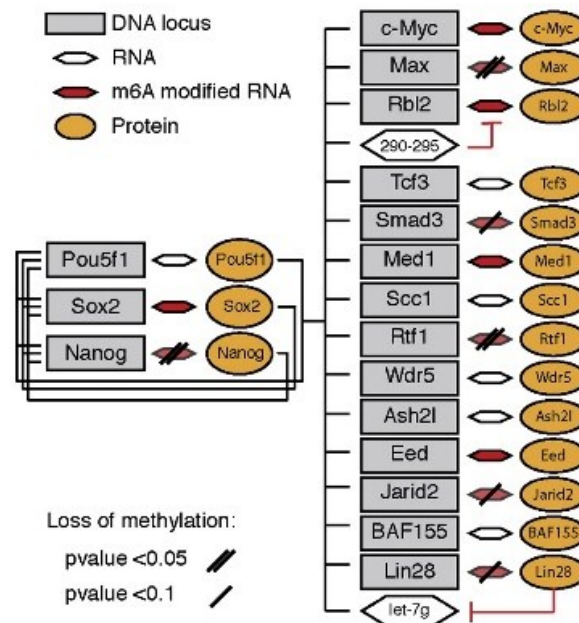


**Fig. 1.9 The mRNA of core stem cell genes contain m6A.**

Model of the genes involved in maintenance of stem cell state (adapted from Young, 2011), representing transcripts with loss of m<sup>6</sup>A modification in *Mettl3*<sup>-/-</sup> cells (Batista et al., 2014). The statistical significance of m6A loss upon *Mettl3*<sup>-/-</sup> is given through the number of strike-throughs. Figure adapted with permission of the rights holder, Elsevier.

## 1.2 Pluripotency and epigenetic reprogramming

Having established connections between RNA 5mC and m6A in pluripotency, development, and epigenetic reprogramming, it is critical to consider these processes fully before performing experiments within these models.

### 1.2.1 Defining pluripotency

Cells can be defined as being totipotent, pluripotent, multipotent or unipotent, depending upon their potential to form all, all embryonic, many, or one differentiated cell type. Pluripotency *in vivo* occurs in the inner cell mass (ICM), pre-implantation epiblast, and early post-implantation

epiblast of mammalian embryos (Boroviak and Nichols, 2014). The extra-embryonic structures, such as the embryonic components of the placenta, are formed from the trophoblast cells which surround the ICM. Inner cell mass pluripotency, of which embryonic stem cells are a parallel, only occurs over a relatively small period of time in the embryo (E3.5–E5.5 in mice and day 5 - 12 humans) before cells become primed to differentiate down a lineage path (Fig. 1.10) (Smith, 2017, Boroviak and Nichols, 2014, Cockburn and Rossant, 2010, Russ et al., 2000). The experimental procedure by which pluripotency can be assessed, is the ability for cells to form all embryonic lineages when injected into an embryo to form a chimaera (Bradley et al., 1984). We thus have a working definition of pluripotency along with a standard metric to experimentally determine it, based upon the ability of a cell to form all embryonic lineages when creating a chimaera. Other pluripotency assays can be used, such as the expression of certain pluripotency markers (such as *KLF4* and *ESRRB*), as well as alkaline phosphatase (AP) staining, however chimaerism remains the gold standard (Boroviak et al., 2015, O'Connor et al., 2008). Naïve and formative pluripotency both fulfil this criteria, and are separated from one another based upon the expression of either naïve (e.g. Kruppel-like factor 2 (*KLF2*)) or early post-implantation factors (e.g. *OTX2*) (Boroviak et al., 2015, Acampora et al., 2016, Kalkan et al., 2017). Functionally, the key distinction between these cells states is that the formative stem cells are responsive to lineage commitment signals, a property referred to as capacitation, while the gene regulatory networks of naïve cells resist such stimuli (Smith, 2017, Kalkan et al., 2017, Turner et al., 2014). This is a reflection of formative cells as a defined intermediate between naïve and primed cells.

Cells which are primed to differentiate in the epiblast, but have crucially not undergone differentiation yet, have been shown to also reflect a form of pluripotency, extending the definition to include up until the point of loss of expression of the pluripotency regulator *OCT4* during somitogenesis (a period of 4-5 days) (Osorno et al., 2012). However, being fated to

differentiate down a particular lineage means that reversion to formative or naïve pluripotency is required to enable the form of all embryonic cell types. As such, primed post-implantation epiblast-derived stem cells (EpiSCs), derived from E5.5-E8.0 embryos, fail to be incorporated into all lineages when inserted into blastocysts or morulas to form chimaeras (Osorno et al., 2012, Masaki et al., 2016). Genetic manipulation to resist apoptosis via BCL overexpression is necessary for EpiSCs to survive insertion at this stage (Masaki et al., 2016). In such a situation, BCL overexpressing EpiSC were found to contribute to all chimaera foetal tissues at E13.5 (Masaki et al., 2016). Chimeras from wild type EpiSCs can be formed through grafting to postimplantation embryos in whole embryo culture (Huang et al., 2012). Together, these experiments show that pluripotency is retained up to and including EpiSC formation, and that the developmental point of chimeric insertion is critical to cell survival and subsequent integration into all embryonic lineages.

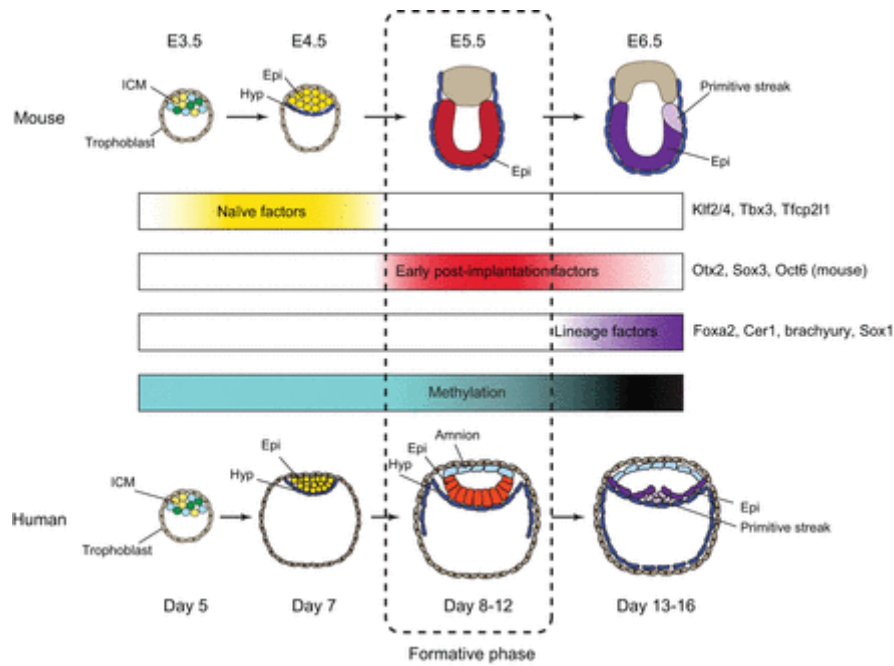


**Fig 1.10 Developmental progression of pluripotency in mouse and human embryos.**

Pluripotent cells begin to emerge in the ICM and segregate to constitute the naïve epiblast. The multi-coloured cells of the ICM indicate mosaic specification of epiblast and hypoblast. After implantation in both mouse (E5) and human (day 8) embryos the epiblast expands as a pseudoepithelial layer overlying the hypoblast (also called the extra-embryonic endoderm), forming a cup-shaped cylinder in mice and a disc in humans. During this period, epiblast cells may remain unpatterned and without molecular specification. Subsequently, epiblast cells become fixed in a columnar epithelium, display regionalised expression of specification factors in response to extra-embryonic signalling centres, and initiate gastrulation. This sequence of events is reflected in transcriptional and epigenetic changes. The distinction between naïve pluripotency and the hypothesised formative phase appears to be acute, whereas the subsequent transition to primed pluripotency is more gradual. Formative and primed phases may be present together at the early stages of gastrulation, particularly in humans. Epi, epiblast; Hyp, hypoblast (Smith, 2017). Figure reproduced with permission of the rights holder, The Company of Biologist Ltd.

The differing pluripotent states which exist between naïve and primed cells are discrete from one another with distinct molecular hallmarks. These are summarised in Table 1.2. Of note, naïve cells are more homogeneous with a greater dependency on mitochondrial respiration than primed cells, while primed cells have undergone X-inactivation, epithelialisation, and apicobasal polarity. Formative cells are notable for their germ cell competence. As outlined above, despite these differences each of these cell states retains pluripotency as assessed by chimera formation and traced by *Oct4* expression.

|                                           | Naïve | Formative     | Primed   |
|-------------------------------------------|-------|---------------|----------|
| Expression of Oct4                        | Yes   | Yes           | Yes      |
| Expression of naïve pluripotency TFs      | Yes   | No            | No*      |
| Expression of early post-implantation TFs | No    | Yes           | Partial‡ |
| Expression of non-ectoderm lineage TFs    | No    | No            | Yes      |
| X inactivation                            | No    | Partial       | Yes      |
| Germ cell competence                      | No    | Yes           | No§      |
| DNA methylation                           | Low   | Intermediate¶ | High     |
| Epithelialisation                         | No    | Partial       | Yes      |
| Apicobasal polarity                       | No    | Partial       | Yes      |
| Homogeneity                               | Yes   | Yes**         | No       |
| Mitochondrial respiration                 | High  | Low**         | Low      |
| Dependence on FGF/Erk signalling          | No    | Low**         | High     |
| Dependence on Nodal/activin signalling    | No    | Low**         | High     |
| Formation of blastocyst chimaeras         | Yes   | Yes**         | No       |
| Formation of post-implantation chimaeras  | No    | Yes**         | Yes      |
| <i>In vitro</i> stem cell derivation      | Yes   | Yes**         | Yes      |

**Table 1.2 Properties of mouse cells in different phases of pluripotency.**

TFs, transcription factors.

\*Nanog is re-expressed in pre-gastrulation posterior epiblast in the mouse egg cylinder and in EpiSCs. In the cynomolgus epiblast Nanog appears to be expressed continuously (Nakamura et al., 2016).

‡ Factors such as Oct6 and Otx2 are upregulated throughout the early postimplantation epiblast but later become restricted to the anterior presumptive neuroectoderm.

§ In human, a subset of primed cells *in vitro* are able to produce primordial germ cell-like cells (Irie et al., 2015; Sasaki et al., 2015).

¶ In formative stem cells, global methylation may be more extensive, as seen for ESCs in serum (Ficz et al., 2013; Habibi et al., 2013; Leitch et al., 2013).

\*\*Predicted properties.

(Smith, 2017). Figure reproduced with permission of the rights holder, The Company of Biologist Ltd.

## 1.2.2 The epigenetic ground state

Waddington's model of an epigenetic landscape, first proposed in 1957, remains an informative analogy for understanding cell fate decisions during development (Fig 1.11) (Waddington, 1957, Goldberg et al., 2007). In this model a totipotent cell resides at the top of a slope which contains multiple ravines. Differentiation is modelled by gravity, which pulls a ball (representing a cell). The multiple final locations at the bottom of the slope represent terminally differentiated cell states. As the cell "rolls" down the hill it will be forced horizontally one way or another at each

undulation in the slope. Upon committing to a direction, some final positions at the bottom of the slope become inaccessible, and so the cell has constrained its fate to a more limited lineage. Which way the cell falls may be purely stochastic, or influenced by lineage determining signals. Through progressive cell fate decisions, the cell eventually reaches the bottom of the slope and resides in a stable, terminally differentiated state. The summit of the slope represents the ground cell state, from which any other cell fate can form. A cell can only be returned to the ground state, which is functionally required at each generation to form a complete multicellular organism, by being driven up the slope against gravity. As the analogy implies, this will not occur stochastically, but rather requires wholesale rewiring of the cell in a process referred to as epigenetic reprogramming.

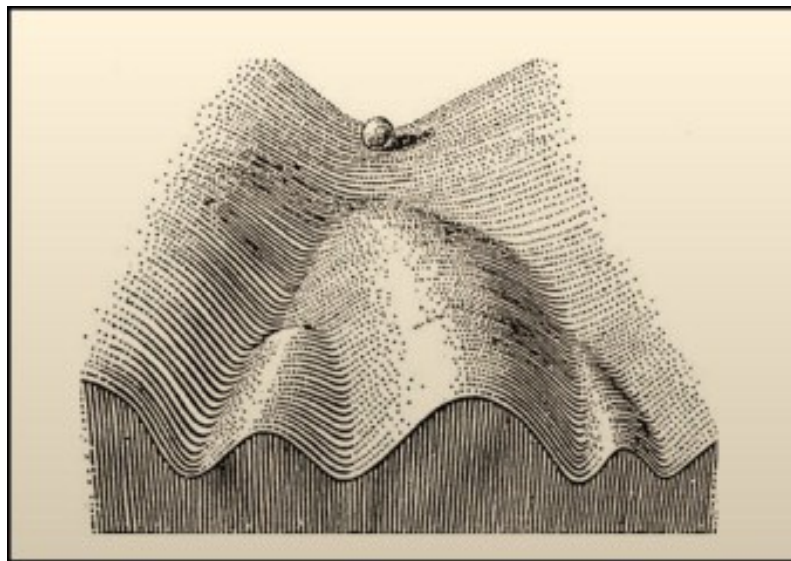


**Fig 1.11 Waddington's Classical Epigenetic Landscape.**

Differentiation is modelled by a ball rolling down an undulating slope. The path of the ball represents progressive cell fate transitions of a totipotent cell. This eventually results in a terminally differentiated cell in one defined position at the bottom of the slope (Waddington, 1957, Goldberg et al., 2007). Figure reproduced with permission of the rights holder, Elsevier.

The Greek prefix epi- identifies epigenetics as being a layer of control that acts “on top of” genetics. In 1942 Waddington first coined the term considering it to be “the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being” (Waddington, 1942). It can thus be considered to involve anything which controls the function of a gene other than the DNA base sequence. To date this has involved such seemingly divergent mechanisms as DNA methylation, histone variants, histone modifications, histone insertions, removal and repositioning, topological chromatin looping, positioning of chromosomes within the nucleus, RNA inheritance, piRNAs, and R-loops (Wallrath and Elgin, 1995, Leitch et al., 2013, Hill, et al., 2014, Nashun et al., 2010, Hajkova et al., 2008, Zhang and Reinberg, 2001, Kuo et al, 1998, Bazett-Jones et al., 1999, Whitehouse et al., 1999, Merckenschlager and Nora, 2016, Finlan et al., 2008, Rassoulzadegan et al., 2006, Aravin et al. 2008, Skourti-Stathaki et al., 2014). This non-exhaustive list does not include RNA modifications, as they have been considered to form the basis of an additional layer of regulation termed epitranscriptomics. Should an RNA modification be shown to feedback to the control of the genomic locus of the transcript it resides in, then this would form an additional area of epigenetics.

Epigenetics can be thought of as transcriptional memory. Indeed, a school of thought in the historical debates over cause and effect in regards to epigenetics and transcription proposed this to be the sole function of epigenetics. The cause/consequence debate hinges upon the temporal order of the epigenetic cis and trans elements and RNA polymerase machinery at any given genomic locus. The clearest indication of epigenetics inducing a novel transcriptional state in a wild type *in vivo* context is transgenerational epigenetic inheritance (Heard and Martienssen, 2014, Radford et al., 2014, Ashe et al., 2012, Belicard et al., 2018). In *C.elegans* the piRNA pathway can establish transcriptional silencing across at least 20 generations independent of the initial silencing signal (Ashe et al., 2012). Transgenerational epigenetic inheritance mires the

ability to define a single ground cell state, as parental effects remain. Regardless, it has long been clear that 5mdC DNA methylation is key for the stable adoption of cell fates, and that the removal of this epigenetic modification from numerous genomic loci is necessary in order to reprogramme a cell to an epigenetic ground state (Hajkova et al., 2002).

### 1.2.3 Mechanisms of epigenetic reprogramming *in vivo*

Gene regulatory networks (GRNs) act to define the state a cell exists in. Within the mammalian life cycle there are two stages at which the existing GRNs undergo their most wholesale disruption, the transitions between haploid and diploid genotypes during fertilisation and gametogenesis. As the zygote will give rise to all embryonic structures and mouse embryonic germ cells have been shown to be equivalent to mESCs from the inner cell mass (ICM) when derived in the same medium conditions, fertilisation and gametogenesis involve the potency of the cells reverting to either a totipotent or pluripotent state respectively (Leitch et al., 2013). DNA 5mdC modifications are crucial for mammalian cells to stably adopt differentiated states. As a result, their erasure is a defining aspect of cellular reprogramming in both the zygote and primordial germ line (Hill et al., 2014; Sheaffer et al., 2014, Messerschmidt et al., 2014). To understand the epigenetic reprogramming that occurs during these processes, DNA 5mdC methylation must be examined further.

In mammals DNA methylation predominantly occurs at CpG sequences, and exists genome-wide (Lister et al., 2009). It acts to silence the expression of methylated regions of the genome forming permanently repressed constitutive heterochromatin (e.g. in centromeres and telomeres) or facultative heterochromatin at regulated sites that might otherwise be euchromatic (e.g. gene promoters) (Keshet et al., 1985, Beyersmann, 2000, Suzuki and Bird, 2008). Cytosine spontaneously deaminates to uracil, which is detected, excised and repaired by DNA repair machinery (Duncan and Miller, 1980). Methylated cytosines deaminate to thymine, resulting in

T:G mismatches which are more likely to be missed by the repair machinery than U:G, causing point mutations and a gradual reduction of the CpG frequency in highly methylated regions of the genome (Poole et al., 2001, Kondrashov, 2003). As a result of this, CpGs are underrepresented throughout the majority of the genome, with the exception of highly enriched hypomethylated CpG islands. CpG islands are associated with ~70% of genes (Suzuki and Bird, 2008). It is at CpG islands that DNA 5mdC can bring about facultative heterochromatin to stably silence genes. By silencing genes associated with alternative cell fates, DNA methylation restricts the lineage of a cell during differentiation to bring about stable differentiated cell fates (Sheaffer et al., 2014). In line with this, conditional knockout of the maintenance DNA methyltransferase *Dnmt1* causes adherent expression of lineage restricted genes through the loss of enhancer methylation (Sheaffer et al., 2014). Typically, CpG island promoters of key genes in the regulation of development remain poised during development, with both repressive and active histone modifications present. This state persists until resolved into either an active or repressed state (Azuara et al., 2006, Mikkelsen et al., 2007, Bernstein et al., 2006). These bivalent genes are enriched for CpG islands and DNA methylation serves to lock previously bistable CpGs in a decisively repressed status (Mantsoki et al., 2015, Smith et al., 2012, Borgel et al., 2010). It logically follows that 5mdC must be erased to regain the genomic plasticity required to reprogramme a cell back to both a pluripotent and epigenetic ground state and reset its dosage compensation to deal with either gaining or losing an allele of each gene.

#### 1.3.3.1 Zygote and early embryo

Upon fertilisation in mammals, the paternal pronucleus undergoes genome-wide DNA demethylation first, with the exception of imprinted genes, intracisternal A particle (IAP) retrotransposons, and centromeric repeats, which mostly escape the process (Fig 1.12) (Lane et al., 2003, Borgel et al., 2010, Rougier et al., 1998, Saitou et al., 2011). Demethylation of the paternal pronucleus occurs rapidly and is substantially depleted by pronuclear stage 3 (PN3),

which represents 7-8 hours post fertilisation (Mayer et al., 2000, Amouroux et al., 2016). Experiments using aphidicolin to block DNA replication do not prevent paternal DNA demethylation (Oswald et al., 2000). This, together with the rapid kinetics of the process, shows that DNA demethylation in this situation does not merely occur through a passive mechanism of dilution during DNA replication, but rather necessitates an active process. It was accepted in the field that 5mdC was converted to 5hmdC by TET3. However, the increase of 5hmdC occurs too late for this to be the case and *Tet3* knockout or chemical inhibition do not prevent demethylation (Hill et al., 2014, Amouroux et al., 2016, Hill et al., 2018). Instead 5hmC acts to protect demethylated sites from *de novo* methylation (Amouroux et al., 2016). The mechanism by which the paternal pronucleus is actively demethylated remains uncertain, although the BER pathway is likely to be involved (Hajkova et al., 2010, He et al., 2011, Santos et al., 2013).

DNA methylation is not the only major epigenetic mark involved in epigenetic reprogramming in the male pronucleus. Paternal DNA is packaged around protamines, which are exchanged with maternal histones within the first hour post-fertilisation (Li et al., 2010). The nucleolus, which acts as an RNA modification hub, develops in tandem from nucleolus precursor bodies to form functional nucleoli capable of zygotic rRNA transcription (Hernandez-Verdun et al., 2010, Fléchon and Kopečný, 1998). This is ahead of the maternal-zygotic transition, in which maternal RNAs are degraded and replaced by the activation of the zygotic transcription at the two cell stage (O'Farrell et al., 2004). The first cell divisions occur rapidly without an increase in total cytoplasmic volume. This permits zygotic transcriptional activation through repressor titration as the nuclei:cytoplasmic ratio increases (Prioleau et al., 1994, Schier, 2007). These additional epigenetic and transcriptional transitions are considered alongside DNA methylation during the experimental work in this thesis.

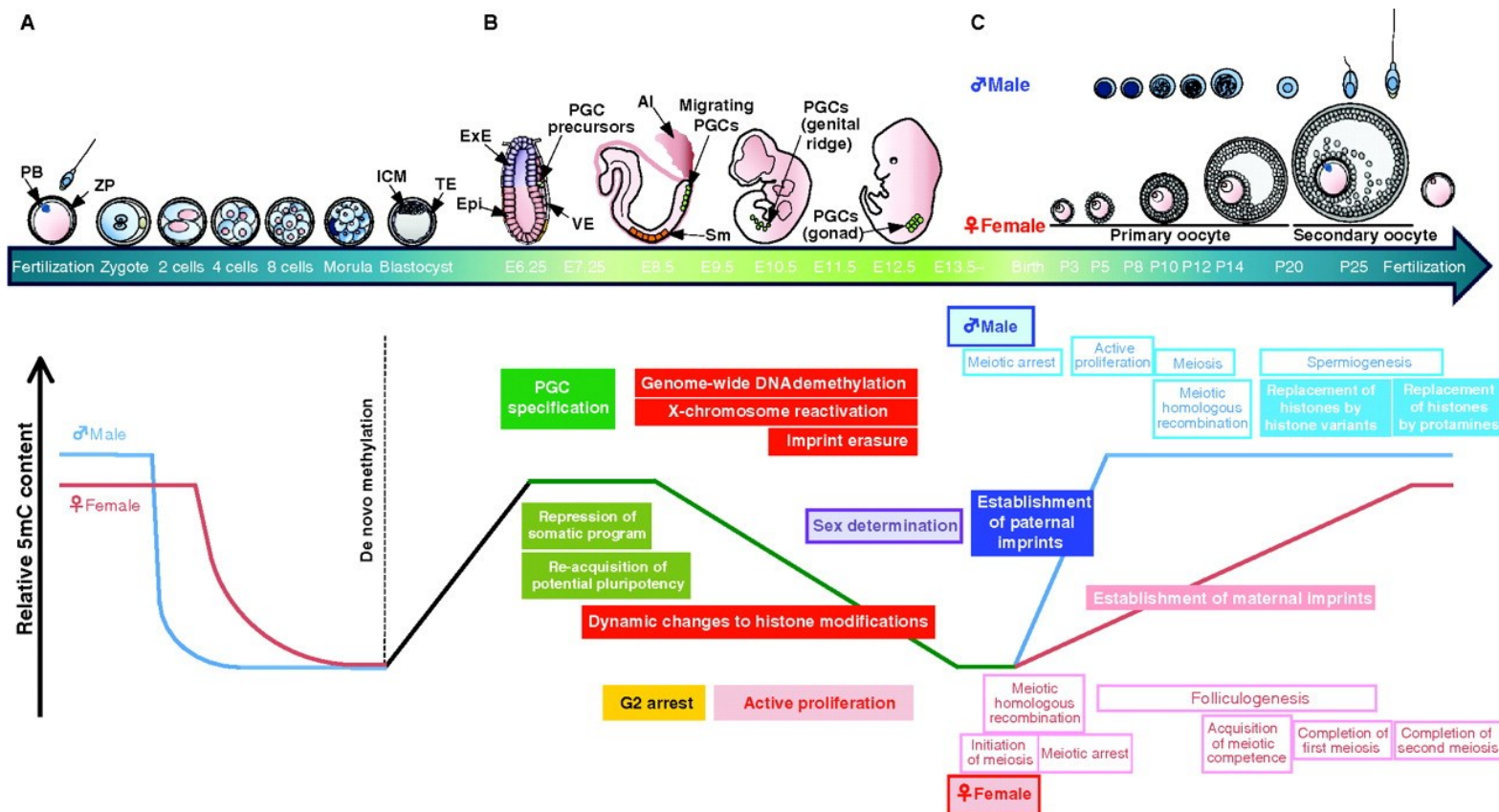


**Fig. 1.12** A schematic of mouse pre-implantation and germ cell development.

(Top) A schematic of pre-implantation and germ-cell development in mice. (A) Pre-implantation development stages; (B) post-implantation embryonic development, following blastocyst implantation at around E4.5; and (C) postnatal germ cell development and maturation. Primordial germ cell (PGC) precursors (E6.25) and PGCs are shown as green circles in embryos from E6.25 to E12.5. (Bottom) Key genetic and epigenetic events are shown that are associated with pre-implantation and germ cell development, together with relative levels of 5-methylcytosine (5mC) at different developmental stages. AI, allantois; Epi, epiblast; ExE, extra-embryonic ectoderm; ICM, inner cell mass; PB, polar body; PGCs, primordial germ cells; Sm, somite; TE, trophoblast; VE, visceral endoderm; ZP, zona pelucida (Saitou et al., 2011). Figure adapted with permission of the rights holder, The Company of Biologist Ltd.

The maternal chromosomes begin demethylation after the paternal ones, through what appears to be a passive loss of 5mdC from the first S-phase of the 1 cell embryo to a minimum around the morula or early blastocyst stage (Fig. 1.12) (Kafri et al., 1992, Monk et al., 1987, Howlett et al., 1991). This passive loss of DNA 5mdC also affects the paternal DNA, resulting in an even lower level of methylation. The extent of CpG demethylation from oocytes to E3.5 embryos is less than would be expected if 5mdC is halved at each cellular division, and so some degree of maintenance of *de novo* methylation must occur (Smallwood et al., 2011). Indeed, imprinted genes, IAP retrotransposons, and centromeric repeats maintain DNA methylation during this period in all chromosomes (Lane et al., 2003, Borgel et al., 2010, Rougier et al., 1998). DNMT1 has been shown to be located outside the nucleus during early embryo development by immunofluorescence (Hirasawa et al., 2008). As DNMT1 acts to maintain methylation patterns during DNA synthesis by depositing 5mdC on the newly synthesised strand of hemimethylated DNA, its exclusion from the nucleus could represent a workable hypothesis for progressive replicative loss of 5mdC. Knockout of *Dnmt1* causes loss of methylation at imprinted genes, indicating that DNMT1 is still active during this stage of development in which it acts upon specific regions of the genome as opposed to all of it (Hirasawa et al., 2008).

While this passive epigenetic reprogramming occurs it is critical to consider the cell fate transitions that are concurrently occurring. The first cell fate transition within a mammal arises in the blastocyst at the segregation of the outer cells of the late morula from the inner cells. The outer cells will form the trophoctoderm in the blastocyst, while the inner cells form the ICM, from which both the epiblast and primitive endoderm form (Rossant and Tam, 2009). All 3 lineages retain pluripotency, but only the epiblast lineage gives rise to embryonic cells. Despite retaining pluripotency, epigenetic changes occur at lineage specific genes associated with the three lineage affiliations (Santos et al., 2010). This has been shown through ChIP-PCR of active (e.g. H3K9ac) and repressive (e.g. H3K27me3) histone modifications at these lineage-associated loci,

with differing replication timings at these sites in the three lineages showing that DNA accessibility is altered (Santos et al., 2010).

The key molecular regulators involved in these early cell fate transitions are OCT4, Caudal Type Homeobox 2 (CDX2), Nanog and GATA-binding factor 6 (GATA6). In the morula OCT4 is uniformly expressed while CDX2, Nanog and GATA6 are all stochastically expressed within the cells. CDX2 then becomes exclusively expressed in the developing trophectoderm through upregulation by apical domain and inhibition from the internal environment (Dietrich and Hiiragi, 2007; Ralston and Rossant, 2008). CDX2 then initiates trophoblast differentiation (Niwa et al., 2005). CDX2 represses OCT4 expression, which consequently becomes exclusively expressed in the ICM. Nanog and GATA6 define the separating epiblast and primitive endoderm lineages (Chazaud et al., 2006; Gerbe et al., 2008). While these genes are stochastically expressed in the morula, from the early blastocyst the master regulator growth receptor bound protein 2 (GRB2) acts to upregulate GATA6 and downregulate Nanog (Chazaud et al., 2006). This generates mutually exclusive expressive patterns that locks in the cell fate decision (Chazaud et al., 2006, Gerbe et al., 2008).

#### 1.3.3.2 Primordial germ cells

Primordial germ cells (PGCs) are specified in the proximal epiblast at ~E6.0 as a result of bone morphogenetic protein (BMP) signalling (Lawson et al., 1999). At this stage the epiblast cells are in the process of losing pluripotency, however the PGC-specified cells retain the pluripotency gene regulatory network (Kurimoto et al., 2008). This is brought about by the upregulation of the key pluripotency regulators B lymphocyte-induced maturation protein-1 (BLIMP1) and PRDM14 in response to BMP signalling (Ginsburg et al 1990, Ohinata et al., 2009, Ohinata et al., 2005, Vincent et al., 2005, Yamaji et al., 2008). From here they migrate to the genital ridge, via the hindgut, in a process that concludes around E11.5 (Anderson et al., 2000, Tam et al.,

1992, Hajkova et al., 2002). Between E8 and E9.5 there is an initial gradual wave of DNA demethylation (Seisenberger et al., 2012). This is considered to be passively brought about by the transient downregulation of DNMT1 and Ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1), through which DNMT1 associates to DNA (Seki et al., 2005, Kagiwada et al., 2012). Upon entering the genital ridge, PGCs are thought to receive cell non-autonomous signalling to undergo epigenetic reprogramming involving a much more rapid DNA demethylation which includes imprinted regions (Hajkova et al., 2002). The nadir of DNA methylation is achieved within one day in both male and female embryos, resulting in ~0.3% of all deoxycytidines being methylated compared to ~4% in somatic tissues (Hajkova et al., 2002, Hill et al., 2018). Even more rapid losses of 5mdC can be observed at specific loci at this time by bisulphite sequencing (Hajkova et al., 2002, Seisenberger et al., 2012, Hill et al., 2018). Such kinetics suggest an active demethylation mechanism. Recently work within the Hajkova lab made substantive advances in unravelling the molecular mechanisms by which this occurs. Much like demethylation of the paternal zygotic pronucleus, 5hmdC is required to maintain demethylated regions by preventing *de novo* methylation (Hill et al., 2018). However, as in the male pronucleus of the zygote, it is unlikely to represent the primary demethylation intermediate due to its relatively uniform and low abundance (Hill et al., 2018). Further, though global 5hmdC levels are stable in PGCs undergoing reprogramming, it gradually decreases at uniquely mapped regions as quantified by AbaSI endonuclease digestion coupled with sequencing (ABA-seq), redistributing to repetitive elements which resist demethylation (Sun et al., 2013, Hill et al., 2018). Given this very rapid turnover kinetics 5hmdC removal would be necessary for this pathway to generate the flux necessary to be the primary demethylation mechanism as had been hypothesised earlier (Hackett et al., 2013).

Alternative proposed demethylation pathways in PGCs include the deaminases activation-induced deaminase (AID) and apolipoprotein B mRNA editing enzyme, catalytic polypeptides

(APOBECs). Both had been thought to act to catalyse the deamination of cytosine or 5-methylcytosine to uridine or thymine respectively, which are subsequently repaired by BER (Morgan et al., 2004). However, both enzymes show reduced activity towards modified deoxycytidines relative to unmodified ones, casting major doubt over the veracity of this mechanism *in vivo* (Nabel et al., 2012). Single-stranded DNA (ssDNA) breaks resulting from AID-induced deamination are necessary for class switch recombination in the Ig locus, providing a further model of possible action in DNA demethylation (Muramatsu et al., 2000, Revy et al., 2000). *Aid* knockout mice PGCs are unable to demethylate to the same extent as wild type PGCs, while somatic cells display wild type methylation levels (Popp et al., 2010). This indicates that AID is necessary, but not sufficient, for complete PGC DNA demethylation. This proposed mechanism remains controversial, as the lack of reproductive defects in *Aid*<sup>-/-</sup> mice put into question the biological relevance AID-dependent component of demethylated sites (Popp et al., 2010). *In vitro* work has shown that AID has a much lower activity for methylated C than unmodified, which puts AID's proposed role in DNA demethylation further in question (Larijani et al., 2005). As with the male pronucleus, one of the few clear details about DNA demethylation in PGCs is that the BER pathway is probably entailed (Hajkova et al., 2010, He et al., 2011).

While DNA demethylation is a defining part of epigenetic reprogramming, there are other key epigenetic events involved in erasing the cell identity. Histone chaperones HIRA and nucleosome assembly protein 1 (NAP-1) are involved in histone variant exchange (Hajkova et al., 2008). Upon entry to the gonad at E11.5 PGCs display a loss of the linker histone H1, along with global reductions in H3K9me3 and H3K27me3 (Hajkova et al., 2008). These changes to histones architecture act synergistically with DNA demethylation to derepress facultative heterochromatin. During reprogramming PGCs undergo substantive metabolic and translational shifts which are characteristic of this transition. The pluripotency gene OCT4 causes the upregulation of the glycolytic enzymes hexokinase, pyruvate kinase, and pyruvate dehydrogenase

kinase (Kim et al., 2015). This is reinforced by wide-reaching translational regulation brought about through miRNAs such as lin28. Lin28 is involved in the regulation of genes associated with glucose metabolism, glycolysis, cellular carbohydrate metabolism, and oxidative metabolism in human ESCs (Polesskaya et al., 2007, Peng et al, 2011, Zhu et al., 2011).

Mouse PGCs and early embryos were used as an *in vivo* model of epigenetic reprogramming in the experimental work performed within this thesis.

#### 1.2.4 Modelling epigenetic reprogramming *in vitro*

The *in vivo* systems outlined above involve low cell numbers from tissue which is difficult to acquire, and so additional experiments were performed in a tissue culture system. Mouse embryonic stem cells cultured in serum-containing media are transcriptionally similar to mouse E10.5 PGC cells by RNA-seq transcriptome analysis (Hill et al., 2018). They also provide a more amenable model to manipulation in regards to both DNA methylation and genetics than *in vivo* mouse models, and can be used to generate large amounts of material. They can form the basis of an *in vitro* reprogramming system if wild type (denoted E14 or J1 based upon the cell line) and *Dnmt1*, *Dnmt3a*, and *Dnmt3b* triple knockout (TKO) ES cells are grown in both serum containing media conditions and 2i. 2i refers to a culturing condition containing two inhibitors, to mitogen-activated protein kinase (MAPK) and glycogen synthase kinase 3 $\beta$  (GSK3 $\beta$ ) respectively, which induce global DNA demethylation (Leitch et al., 2013; Ying et al., 2008; Ficz et al., 2013; Habibi et al., 2013). It had been believed that genomic imprints are retained in ES cells cultured in 2i conditions, however it has been recently detailed that prolonged ES culture in 2i conditions can cause imprint erasure (Leitch et al., 2013; Choi et al., 2017; Yagi et al., 2017). Cells grown in 2i display reduced expression of the *de novo* DNA methyltransferases, and therefore this has been hypothesised to be at least part of the mechanism by which this reduction in global DNA methylation is brought about (Leitch et al., 2013). TKO

cells lack the canonical de novo and maintenance DNA methyltransferases and display no detectable DNA methylation above background noise regardless of media condition (Tsumura et al., 2006, Ficz et al., 2011, Williams et al., 2011). Through comparison of wild type ES cells grown in serum, wild type ES cells grown in 2i, and TKO cells grown in either condition, environments of descending DNA methylation from ~4% to ~2% to 0%, as determined by the ratio of 5-methyldeoxycytosine to deoxyguanosine (5mdC/dG), can be assessed (Leitch et al., 2013, Tsumura et al., 2006). This represents a more extreme example of DNA demethylation than is observed in wild type in vivo contexts, and so represented a potentially useful model for screening for 5mdC/5mC or DNMT/RNMT interplay.

### 1.3 Key Hypotheses

- In the establishment of DNA 5mdC as a mechanism to regulate genome utilisation, RNA 5mC has become connected to genomic control and developmental processes. This is facilitated through an incomplete separation of the molecular machineries associated to these two evolutionarily-linked nucleoside modifications
- RNA 5mC and 5hmC have biologically relevant functions, some of which are more relevant in, or are specific to, developmental processes such as epigenetic reprogramming
- RNA 5mC and 5hmC levels change globally and locally during the time-course of epigenetic reprogramming in response to the global erasure of DNA methylation as an epigenetic regulatory mechanism
- RNA 5mC exhibits a functional interplay with 5mdC
- 6-methyladenine is involved in the cell fate transitions, and so is expected to alter in its abundance during epigenetic reprogramming

# Chapter 2: Materials and methods

## 2.1 Mice

### 2.1.1 Animal husbandry

All animal work was performed under a UK Home Office approved personal licence and project licence, as well as taking place within a facility with a UK Home Office approved establishment licence. C57BL/6J (Charles River, C57BL/6J) and B6CBAF1/J (The Jackson Lab, B6CBAF1/J) mice were purchased externally and bred with homozygous GOF18 $\Delta$ PE-EGFP mice (Yoshimizu et al., 1999). The PGCs of the resulting embryos expressed GFP under the OCT4 promoter, which facilitated downstream experimental work.

### 2.1.2 Genotyping

#### 2.1.2.1 DNA extraction

Material for genotyping was taken by either ear sniping for pups and adults, or removing a paw from embryos while dissecting. DNA was extracted in 50  $\mu$ l of genotyping digestion buffer (50mM KCL (AnalaR NORMAPUR, 26764.260), 10 mM Tris-HCl (Sigma, T5941) pH 8.3, 0.45% (v/v) Tween-20 (Sigma, P7949), 0.1 mg/ml gelatin (Sigma, G1393) in H<sub>2</sub>O), with the fresh addition of proteinase K (NEB, P8107S) to 200  $\mu$ g/ml, by incubation overnight at 55 °C.

#### 2.1.2.2 Phenol:chloroform:isoamyl alcohol extraction

An equal volume of Phenol:chloroform:isoamyl alcohol (Invitrogen, 15593.031) was added to the digested samples which were mixed by pipetting. Centrifugation was performed at 12,000  $\times$ g for 5 minutes before the aqueous phase was removed and 5  $\mu$ l of 3M Na Acetate (AnalaR NORMAPUR, 27653.260) pH 5.2 added. DNA was precipitated by making the solution to 70% (v/v) ethanol and incubating at -20 °C for 30 minutes. The precipitated DNA was centrifuged at

16,000 xg for 15 minutes, washed in 70% (v/v) ethanol before centrifuging again. The supernatant was removed and the samples dried using a speed vacuum (Thermo, Savant SPD121P). Samples were finally resuspended in H<sub>2</sub>O and DNA concentrations quantified by NanoDrop.

### 2.1.2.3 PCR

GOF18ΔPE-EGFP genotyping was performed by RT-qPCR to enable homozygous and heterozygous individuals to be identified by insertion copy number (see section 2.12). To determine the sex of embryos, PCR of the Y chromosome gene SRY was performed using RedTaq Ready Mix with 1 µl of extracted DNA and forward and reverse primers at a final concentration of 400 nM in 20 µl. The PCR conditions used are outlined in Table 2.1. The sequence of all primers used throughout this body of work are collated in Table 2.8

| Step             | Temperature [°C] | Duration   |
|------------------|------------------|------------|
| 1                | 95               | 5 minutes  |
| 2                | 95               | 45 seconds |
| 3                | 55               | 45 seconds |
| 4                | 72               | 45 seconds |
| Go to step 2 x34 |                  |            |
|                  | 72               | 10 minutes |
|                  | 4                | ∞          |

**Table 2.1 SRY PCR conditions.**

### 2.1.2.4 DNA gel electrophoresis

Gels were created by first heating 3% (w/v) agarose in 0.5x TBE in a microwave until solvation before cooling, subsequent addition of 3 µl SYB Safe DNA Gel Stain (Invitrogen, S33102), and

setting in casts. Gels were loaded with samples and a 100 b.p. DNA ladder (NEB, N3231S) and run at 100v for 30 minutes – 1 hour in 0.5x TBE before imaging in a gel doc (Syngene).

### 2.1.3 Dissections

#### 2.1.3.1 Genital Ridge

Female mice were euthanised at 10.5-15.5 days post coitum (d.p.c.) by cervical dislocation with cessation of heartbeat as a confirmation. This was assessed by decapitation. The cadaver was dissected open and, if pregnant, the uterus containing the embryos was removed and transferred to PBS on ice. Dissections were performed in filtered 10% (v/v) FCS (Gibco, 10270-106) in PBS under a dissecting microscope. The pregnancy sack and amniotic membrane were ruptured, and embryos culled by severing the umbilical cord before swift decapitation. Embryos aged E11.5 onwards had their anterior abdominal organs removed to expose the genital ridges which were dissected out for downstream applications. Germ cell migration to the genital ridge only concludes around E11.5, and so for embryos aged E10.5 the whole lower trunk was dissected out for downstream applications (Hajkova et al., 2002).

#### 2.1.3.2 Oocytes

Cumulus-oocyte complexes were dissected by Dr. Buhe Nashun as described previously (Nashun et al., 2010, Inoue et al., 2010).

## 2.2 Cell culture

### 2.2.1 Embryonic stem cells

The cell lines used are outlined in Table 2.2. E14, J1, TKO, TET TTKO, EpiSC Murine ES cells were cultured in 15% (v/v) serum or 2i media conditions. Serum media conditions consisted of GMEM (Gibco, 21710-025) supplemented with 15% (v/v) FBS (Gibco, 10270-106) 1x NEAA

(Gibco, 11140-050), 0.1 mM  $\beta$ -mercaptomethanol (Gibco, 31350-010) 1mM sodium pyruvate (Gibco, 11360-039), and 2 mM glutamine (Gibco, 25030-024). GMEM also forms the base for 2i media, being used at a 1:1 ratio with Neurobasal (Gibco, 21103-049). To this 0.5x N2 supplement (Gibco, 17502-048), 0.5x B27 supplement with VitA (Gibco, 17504-044), 0.1 mM  $\beta$ -mercaptomethanol, and 2 mM glutamine were added. Media was filtered and to each media condition 1500 U of ESGRO mLIF (Millipore, ESG1107) was added fresh to each 50ml aliquot. Additionally, 1uM CHIR99021 (Cell Guidance Systems, SM13-10) and 3uM PDO325901 (Stemgent, 04-0006) were added to aliquots of 2i.

Cells cultured in serum media were grown in plates coated with 0.1% (v/v) gelatin (Sigma, G1393) while those cultured in 2i were grown in laminin-coated (Sigma, L2020) plates. Splitting occurred every 2-5 days at a ratio of 1:10-30. Whenever culturing conditions were altered, cells were deemed to have acclimatised after a minimum of 5 passages as previously defined (Leitch et al., 2013).

| Cell line | Genotype                                                                 | Genetic background | Source                                                   |
|-----------|--------------------------------------------------------------------------|--------------------|----------------------------------------------------------|
| E14       | Hypoxanthine phosphoribosyl-transferase (HPRT) <sup>-/-</sup>            | 129/Ola            | Dr. Rachel Amouroux, Tsuda et al., 1997                  |
| J1        | Wild type                                                                | 129/SvJ            | Dr. Rachel Amouroux                                      |
| TKO       | DNMT1 <sup>-/-</sup> , DNMT3A <sup>-/-</sup> , and DNMT3B <sup>-/-</sup> | 129/SvJ            | Dr. Rachel Amouroux, Tsumura et al., 2006                |
| TET TTKO  | TET1 <sup>-/-</sup> , TET2 <sup>-/-</sup> , and TET3 <sup>-/-</sup>      | 129/SvJ            | Dr. Cristina Requena Torres                              |
| EpiSC     | Wild type                                                                | 129                | Prof. Austin Smith, Dr. Carla Mulas and Dr. Harry Leitch |

**Table 2.2 Stem cell lines.**

### 2.2.2 Epiblast stem cells

The cell lines used are outlined in Table 2.2. EpiSC media was created using 1:1 GMEM with Neurobasal to which 0.5x N2 supplement (Gibco, 17502-048), 0.5x B27 supplement without VitA (Gibco, 12587-010), 0.05 mM  $\beta$ -mercaptomethanol, and 2 mM glutamine were added. Media was filtered and 20 ng/ml Activin A (PeproTech, 120-14E-10) and 12 ng/ml bFGF (fibroblast growth factor-basic) (ThermoFisher, 13256029) were added. Splitting occurred every 2-3 days at a ratio of 1:12-15 by manual dissociation. Colonies were plated into tissue culture dishes which had been pre-incubated with FCS which was aspirated before the introduction of cells. Whenever culturing conditions were altered, such as in the generation of TKO EpiLSCs from TKO serum ESCs, cells were deemed to have acclimatised after a minimum of 5 passages. As such, all cell lines created from media transfer were self-renewing.

### 2.2.3 *In vitro* fertilisation and early embryo culture

*In vitro* fertilisation (IVF) was carried out by Dr. Buhe Nashun as described previously (Nashun et al., 2015, Nashun et al., 2010). Briefly, mice were super-ovulated using 5 IU pregnant mare's serum gonadotropin (PMSG) injected intraperitoneally into 3-week old mice. After 48 hours, 5 IU of human chorionic gonadotropin (hCG) was administered intraperitoneally. Cumulus-oocyte complexes were dissected and transferred to human tubal fluid (HTF) medium which had been supplemented with 10 mg/ml BSA (Sigma-Aldrich). Spermatozoa were obtained from the caudal epididymides of adult male mice (Charles River) before being capacitated in HTF medium for 2 hours and then used for insemination. Fertilised oocytes were then washed in HTF prior to culturing in potassium simplex optimisation medium (KSOM). Embryos were cultured until collection at the 1 cell, 2 cell, 4 cell, morula or blastocyst stage.

## 2.3 Florescence-activated cell sorting (FACS)

FACS was performed based upon a previously published methodology (Hajkova et al., 2008). Briefly, E10.5 embryo trunks or E11.5 – E15.5 genital ridges were dissected as indicated above and transferred into filtered 1% (w/v) BSA (Sigma, A3311), 5 mM EDTA (Sigma, E9884), in PBS on ice. Material was centrifuged at 400 xg, resuspended in 200 µl 0.05% trypsin-EDTA (1x) (Gibco, 25300-054) and incubated at 37 °C for 5 minutes. Digestion was quenched with equal volumes of serum media and cells dissociated by pipetting up and down 20x. The cells were centrifuged as before, washed and further dissociated by resuspension in 100 µl 1% (w/v) BSA, 5 mM EDTA in PBS, and centrifuged. The cells were finally resuspended in 100 µl 1% BSA (w/v), 5 mM EDTA in PBS before being passed through a Corning Falcon test tube with cell strainer snap cap (Corning 352235) and then washed through in a further 300 µl. Separation of GFP positive and negative cells was performed using either an Aria Fusion (BD Bioscience) or Aria II (BD Bioscience) flow cytometer with cells sorted into 1% BSA (w/v), 5 mM EDTA in PBS before snap freezing.

## 2.4 Western blotting

### 2.4.1 Protein extraction and quantification

Cells were lysed in RIPA buffer (150 mM NaCl, 1% (v/v) NP-40, 0.5% (v/v) sodium deoxycholate (DOC), 50 mM Tris HCl pH 7.5, 5 mM EDTA pH 8) to which 1x Complete mini, EDTA-Free protease inhibitor (Sigma, 11836170001) and 1x PhosStop (Sigma, 4906837001) were added fresh to aliquots. 5 cycles of sonication at 30 seconds on and 30 seconds off physically lysed the cells in RIPA buffer with the resulting solution centrifuged at 14,000 xg for 15 minutes. The supernatant was taken for total protein concentration quantification using the Bradford assay with a BSA standard curve. Briefly, a standard curve of BSA ranging from 2 µg – 16 µg was generated

and mixed with Bio-Rad protein assay added to 20% (v/v) up to 200 µl in a flat-bottomed well of a 96 well plate. 1 µl of undiluted samples, as well as dilutions ranging to 1:100, were used in technical triplicate across different wells. Absorbance was determined using a SpectraMax 340PC Microplate Reader at 595 nm and used to quantify the protein concentration of the samples.

#### 2.4.2 Casting and running gel

Gels were cast with a stacking gel above the running gel to narrow protein bands. Running gel mix consisted of; 375 mM Tris pH 8.8 0.1% (v/v) sodium dodecyl sulphate (SDS), 0.1% (w/v) ammonium persulphate (APS) and 0.1% (v/v) N,N,N'-tetramethylethylenediamine (TEMED), with acrylamide and H<sub>2</sub>O added appropriately to give a 10% (v/v) -12.5% (v/v) acrylamide gel. Once the running gel was set, stacking gel was created using 125 mM Tris pH 6.5, 0.1% (v/v) SDS, 0.1% APS (w/v) and 0.16% (v/v) TEMED, with acrylamide and H<sub>2</sub>O added appropriately to give a 5% (v/v) acrylamide gel.

2x Laemmli buffer (120 mM Tris pH 6.8, 20% Glycerol, 4% SDS, 0.02% bromophenol blue) was added to equal amounts of total protein, typically 20 µg, followed by heating at 95 °C for 5 minutes. Samples, along with Color Prestained Protein Standard, Broad Range (NEB, P7712S), were loaded into the previously cast gels and run in 1x TGS buffer (Bio-Rad, 161-0772) at 50 V until out of the stacking gel and then at 100 V subsequently.

After electrophoresis, the proteins were transferred from the gel to an activated PVDF membrane. A Bio-Rad Mini-Gel system was used to create a gel sandwich which was submerged in 4 °C transfer buffer (1x TG buffer, 5% (v/v) ethanol, 0.2% (v/v) SDS) and run at 100 V for 1 hour at 4 °C with the addition of an ice pack. The membrane was then blocked in 5% (w/v) non-fat milk in TBS-T (0.05% Tween-20 (Sigma, P7949)) for an hour. Primary antibodies were diluted in 2% (w/v) non-fat milk in TBS-T and the membranes were incubated in this solution

overnight at 4 °C. Membranes were washed 4x ≥ 10 minutes in TBST, incubated with HRP-conjugated secondary antibody for 1 hour and washed again for 4x ≥ 10 minutes in TBST. Luminata Forte Western HRP substrate (Millipore, ELLUF0100) was used for detection on a GE ImageQuant LAS 4000 mini. The antibodies used are detailed in Table 2.3. The images generated were analysed and signals quantified using FIJI software.

Where performed, stripping of primary and secondary antibodies was achieved using stripping buffer (0.2M glycine, 3.5 mM SDS, 1% (v/v) tween-20, pH 2.2) in which membranes were washed for 3x 5 minutes. Membranes were washed and equilibrated with TBST, reblocked, and stripping verified by 2° antibody incubation and imaging as described above. If the signal was completely ablated, then membranes were incubated with the next 1° antibody of interest and from there imaging was performed as described above.

| Target                                  | Host   | Company                    | Reference              | Application                               | Dot Blot dilution | WB dilution | IF dilution stem cells | IF dilution PGCs |
|-----------------------------------------|--------|----------------------------|------------------------|-------------------------------------------|-------------------|-------------|------------------------|------------------|
| 5mc                                     | Mouse  | Millipore                  | MABE146                | Primary antibody                          | 1:1000            |             |                        |                  |
| 5mC                                     | Rabbit | Abcam                      | RM231                  | Primary antibody                          | 1:10000           |             |                        |                  |
| 5mC                                     | Rabbit | Cell Signalling Technology | D3S2Z                  | Primary antibody                          | 1:500             |             |                        |                  |
| 5mC                                     | Mouse  | Active Motif               | Abflex, 91187          | Primary antibody                          | 1:1000            |             |                        |                  |
| 5hmC                                    | Rabbit | Active Motif               | Active Motif, 39791    | Primary antibody                          | 1:2500            |             |                        |                  |
| TRDMT1                                  | Mouse  | GeneTex                    | 102B1259.2             | Primary antibody                          |                   |             | 1:50                   |                  |
| TRDMT1                                  | Mouse  | Thermo                     | MA5-16197 (102B1259.2) | Primary antibody                          |                   | 1:250       |                        |                  |
| NOP2                                    | Rabbit | GeneTex                    | GTX105945              | Primary antibody                          |                   | 1:500       | 1:100                  | 1:50             |
| NSUN2                                   | Rabbit | Proteintech                | 20854-1-AP             | Primary antibody                          |                   | 1:1000      | 1:500                  | 1:25             |
| NSUN3                                   | Rabbit | Abcam                      | ab170188               | Primary antibody                          |                   | 1:250       | 1:50                   |                  |
| NSUN5                                   | Rabbit | Proteintech                | 15449-1-AP             | Primary antibody                          |                   | 1:500       | 1:50                   | 1:25             |
| NSUN6                                   | Rabbit | Proteintech                | 17240-1-AP             | Primary antibody                          |                   | 1:1000      | 1:50                   |                  |
| H3 Pan                                  | Rabbit | Millipore                  | 07-690                 | Primary antibody                          |                   | 1:2000      |                        |                  |
| Oct4                                    | Mouse  | BD Biosciences             | 611202                 | Primary antibody                          |                   |             | 1:200                  |                  |
| GFP                                     | Goat   | Abcam                      | Ab5450                 | Primary antibody                          |                   |             |                        | 1:1000           |
| Mouse IgG                               | Donkey | Santa Cruz                 | SC-2096                | HRP-conjugated secondary antibody         | 1:5000            | 1:2000      |                        |                  |
| Rabbit IgG                              | Donkey | Santa Cruz                 | SC-2077                | HRP-conjugated secondary antibody         | 1:5000            | 1:5000      |                        |                  |
| Goat IgG                                | Donkey | Santa Cruz                 | SC-2056                | HRP-conjugated secondary antibody         |                   | 1:5000      |                        |                  |
| Donkey $\alpha$ -mouse Alexa Fluor 488  | Donkey | Invitrogen                 | A21202                 | Alexa Fluor-conjugated secondary antibody |                   |             | 1:300                  |                  |
| Donkey $\alpha$ -rabbit Alexa Fluor 568 | Donkey | Life Technologies          | A10042                 | Alexa Fluor-conjugated secondary antibody |                   |             | 1:300                  |                  |
| Donkey $\alpha$ -goat Alexa Fluor 488   | Donkey | Invitrogen                 | A11055                 | Alexa Fluor-conjugated secondary antibody |                   |             |                        | 1:300            |
| Donkey $\alpha$ -mouse Alexa Fluor 568  | Donkey | Life Technologies          | A10037                 | Alexa Fluor-conjugated secondary antibody |                   |             |                        | 1:300            |
| Donkey $\alpha$ -rabbit Alexa Fluor 647 | Donkey | Life Technologies          | A31573                 | Alexa Fluor-conjugated secondary antibody |                   |             |                        | 1:300            |

**Table 2.3 Antibodies used for dot blotting, western blotting, and immunostaining.**

## 2.5 Immunostaining

### 2.5.1 Embryonic stem cells

Cells were grown on gelatin coated coverslips in their respective media conditions with the addition of 50 U/mL Penicillin-Streptomycin (Gibco, 15070-063). Upon reaching confluency the coverslips were removed and washed in PBS before fixation in 4% (v/v) paraformaldehyde (PFA) (Thermo, 28908) for 15 minutes. Permeabilisation and blocking was performed using blocking solution (1% (w/v) bovine serum albumin (BSA) (Sigma, A3311) and 0.1% (v/v) Triton X-100 (Sigma, T8787)) in PBS for 45 minutes. The coverslips were incubated overnight with primary antibodies in blocking solution in a humidified chamber. The following day the coverslips were washed 3 times in blocking solution and then incubated with secondary antibodies conjugated to Alexa dyes in the dark for an hour at room temperature. The coverslips were washed once in blocking solution and then twice in blocking solution lacking Triton X-100. DAPI was applied at 1 ng/ $\mu$ l for 10 minutes before a final wash step and mounting in Vectashield (Vector laboratories, H-1000) with slides sealed using nail polish. The slides were then imaged on a Leica SP5 confocal microscope. The antibodies used are detailed in Table 2.3. The images generated were analysed using FIJI software.

### 2.5.2 Genital Ridges

Mice were mated and sacrificed at the appropriate time point during pregnancy. Embryos were removed before their genital ridges were dissected out and fixed with 4% (v/v) PFA in PBS for 30 minutes at 4°C. Three wash steps were performed followed by overnight equilibration in 30% (v/v) sucrose at 4°C. The tissue was then embedded in OCT (VWR Chemicals, 361603E) and snap frozen in liquid nitrogen. Tissues were cut into 10  $\mu$ m thick sections using a Leica CM

1950 cryostat and positioned upon polylysine slides (Thermo Scientific, J2800AMNZ). These slides were then frozen at -80 °C.

Slides were equilibrated to room temperature, post-fixed with 2% (v/v) PFA for 3 minutes and washed for 3x 10 minutes with PBS. Antigen presentation was performed by incubating the slides in 10 mM sodium citrate (pH 6.0) at -80 °C for 10 minutes. Slides were quenched in PBS prior to blocking and permeabilisation in blocking buffer (PBS with 1% (w/v) BSA and 0.1% (v/v) Triton TX-100) at room temperature for 1 hour. Overnight incubation was performed at 4 °C with primary antibodies diluted in blocking buffer in a humidified chamber. The remainder of the protocol was performed as described above for immunostaining of stem cells.

### 2.5.3 PGC single cells

Mice were mated and sacrificed at the appropriate time point during pregnancy. Embryos were removed before their genital ridges were dissected out and dissociated in TrpLE Express (Gibco, 12604-013) for 20 minutes at 37 °C before pipetting up and down and quenching in serum media. At the same time slides were incubated with poly-L-lysine solution (sigma, P4707) for 20 minutes. Cells were resuspended in 200 µl serum media, placed on poly-L-lysine-treated slide and incubated for 20 minutes at 37 °C. Subsequently cells were washed in PBS before fixing in 4% (v/v) PFA in PBS for 15 minutes. Three further PBS wash steps were performed before blocking and permeabilisation in blocking buffer (PBS with 1% (w/v) BSA and 0.1% (v/v) Triton TX-100) for ≥ 30 minutes. Overnight incubation was performed at 4 °C with primary antibodies diluted in blocking buffer in a humidified chamber. The remainder of the protocol was performed as described above for immunostaining of stem cells (section 2.5.1).

### 2.5.4 Early embryo

At the appropriate time points, early embryos were adhered to poly-L-lysine-treated slides and stained as outlined for PGC single cells.

## 2.6 RNA extraction

### 2.6.1 QIAGEN RNeasy kit

RNA was extracted using the QIAGEN RNeasy kit (74104) as per protocol. RNA concentrations were quantified by NanoDrop.

### 2.6.2 Zymo Research Duet DNA/RNA MiniPrep kit

RNA was extracted using the Zymo Research Duet DNA/RNA MiniPrep kit (D7001) as per protocol. RNA concentrations were quantified by NanoDrop.

### 2.6.3 TRIzol

Cells were lysed in TRIzol (Ambion, 15596026) by vortexing or passing through a needle depending on cell number. Samples were incubated for 5 minutes at room temperature to ensure the complete dissociation of nucleoprotein complexes before 20% (v/v) chloroform was added. The solution was mixed by vortexing followed by incubation for 5 minutes at room temperature and centrifugation at 12,000 xg for 5 minutes at 4 °C to separate the aqueous and organic phases. The aqueous phase was removed and precipitated by the addition of 0.5 ml isopropanol per 1 ml TRIzol and incubation at -80 °C for 1 hour. The precipitate was pelleted by centrifugation at 12,000 xg for 10 minutes at 4 °C prior to washing in 75% (v/v) ethanol. The samples were spun as before with residual wash buffer removed using a speed vacuum (Savant, SPD121P) and redissolved in Diethyl pyrocarbonate (DEPC) (VWR, A0881.0020)-treated H<sub>2</sub>O. RNA concentrations were quantified by NanoDrop.

## 2.7 RNA quantification and quality assessment

### 2.7.1 Bioanalyzer

Depending upon the amount and length of RNA samples in question either the Agilent RNA 6000 Nano Kit (Agilent, 5067-1511), Agilent RNA 6000 Pico Kit (Agilent, 5067-1513), or Agilent Small RNA kit (Agilent, 5067-1548) were used, all as per protocol.

### 2.7.2 RNA gel electrophoresis

RNA samples were mixed 1:1 (v/v) with RNA Loading Dye (2x) (NEB, B0363A) with ethidium bromide (Sigma, E1510) added to a final concentration of 100 ng/μl. Samples were then heated at 65 °C for 15 minutes before cooling and loading, along with a ssRNA ladder (NEB, N0362S) into a Lonza Reliant Gel System precast gel (Lonza, 54922), which was run in DEPC-treated 0.5x TBE. The gel was imaged using a gel doc. To enhance the signal from single-stranded RNA, the gels were agitated in 1:10,000 SYBR Gold Nucleic Acid Gel Stain (Invitrogen, S11494) in DEPC-treated 0.5x TBE before being reimaged.

## 2.8 Nucleotide digestion

### 2.8.1 DNase treatment

RNase-free DNase set (Qiagen, 79254) was used as per the protocol outlined for QIAGEN RNeasy kit (74104) for both on-column and in-solution digestions, with adjustment of the enzyme incubation time to 20 minutes. DNase 1 (Sigma, 4716728001) and TURBO DNA-free kit (Invitrogen, AM1907) were also trialled as DNases as per protocols.

### 2.8.2 RNase treatment

Purelink RNase A (Invitrogen, 12091-021) was used to digest RNA samples by the addition of 160 µg of RNase A followed by incubation at 37 °C for 1 hour.

### 2.8.3 RNA clean up

Following in-solution nucleotide digestion, the RNA Clean & Concentrator kit (Zymo, R1013) was used as per protocol. Through nuclease treatment and clean-up, the source of signals detected in downstream dot blotting experiments could be elucidated.

## 2.9 RNA dot blot

### 2.9.1 Urea and formamide/formaldehyde methods

RNA was denatured in equal volume urea solution (8M urea and 10% (w/v) sucrose in 1x TBE) or 3x RNA incubation solution (65.0% (v/v) formamide (Acros organics, 327235000), 7.77 % formaldehyde (w/v) (Sigma, F8775), 1.33x MOPS) and incubated at 65 °C for 15 or 5 minutes respectively before cooling on ice. Samples denatured in RNA incubation solution were mixed with equal volumes of 20x SSC. Serial dilutions were created and spotted onto a Hybond-N<sup>+</sup> nylon transfer membrane (GE Healthcare, RPN203B), left to dry for 15 minutes, and then ultraviolet-crosslinked at 1200 Jm<sup>-2</sup>. Later experiments were performed with a 30-minute baking step prior to UV crosslinking. Membranes were then blocked in 5% (w/v) non-fat milk in PBS-T (0.1% (v/v) Tween-20) for an hour. Primary antibodies were diluted in 1% (w/v) non-fat milk in PBS-T and the membranes were incubated in this solution overnight. Membranes were washed 4x ≥ 10 minutes in PBST, incubated with HRP-conjugated secondary antibody for 1 hour and washed again for 4x ≥ 10 minutes in PBST. Luminata Forte Western HRP substrate (ELLUF0100, Millipore) was used for detection on a GE ImageQuant LAS 4000 mini.

Methylene blue stain (0.02% (w/v) methylene blue in 0.3M sodium acetate (pH 5.2)) was added for 10 minutes; washed 3x 10 seconds in milli-Q H<sub>2</sub>O and imaged on a GE ImageQuant LAS 4000 mini. Antibody dilutions are given in Table 2.3. The images generated were analysed using FIJI software.

### 2.9.2 NaOH method

The method used was previously described in Miao et al., 2016, varying only in UV crosslinking to a Hybond-N<sup>+</sup> nylon membrane being performed instead of baking to a nitrocellulose membrane. Briefly RNA samples were digested with equal volumes of 2M NaOH (sigma, S8045) in DEPC-treated H<sub>2</sub>O for 30 minutes at 4 °C. Samples were then dotted onto a Hybond-N<sup>+</sup> nylon membrane, air-dried at room temperature for 15 minutes and then baked for 30 minutes at 80 °C prior to UV crosslinking at 1200 Jm<sup>-2</sup>. From here the remainder of the protocol followed as outlined above for the urea and formamide/formaldehyde methods (section 2.9.1).

## 2.10 RNA subtype selection

### 2.10.1 Poly(A) enrichment

The Dynabead mRNA Purification Kit (Invitrogen, 61006) was used as per protocol.

### 2.10.2 Small RNA enrichment

#### 2.10.2.1 Purelink miRNA isolation kit

The Purelink miRNA Isolation Kit (Invitrogen, K157001) was used as per protocol, with binding buffer optimised. 300 µl binding buffer was added to TRIzol-extracted total RNA samples followed by 300 µl 70% (v/v) ethanol. This mixture was transferred to a spin cartridge and

centrifuged at 12,000 xg for 1 minute. To the flow-through 700 µl 100% ethanol was added before transfer to a new spin cartridge and centrifugation at 12,000 xg for 1 minute. The column was washed twice with 500 µl wash buffer with centrifugation performed as above before a final spin at maximum speed for 3 minutes to remove any remaining wash buffer. The columns were moved to recovery tubes before 50 µl DEPC-treated H<sub>2</sub>O was applied. After incubation for a minute centrifugation at maximum speed for 1 minute eluted the small RNA.

#### 2.10.2.2 Gel extraction

An RNA gel was run as outlined earlier. A Dark Reader DR88X Transilluminator (Clare Chemical Research) and scalpel were used to extract low base pair bands which were further diced using the scalpel before addition of 400 µl gel elution buffer (20 mM Tris-HCl pH 7.5, 0.25 M sodium acetate, 1 mM EDTA, 0.25% (v/v) SDS). Gel fragments were frozen on dry ice for 15 minutes and incubated overnight at room temperature.

The following day the samples were centrifuged at maximum speed for 10 minutes with the RNA-containing supernatant taken for phenol:chloroform:isoamyl alcohol extraction by the addition of 800 µl. After vigorous shaking the sample was centrifuged at 12,000 xg for 5 minutes with 1 ml isopropanol added to the supernatant. RNA precipitation occurred at -80 °C for 1 hour before centrifugation at maximum speed for 10 minutes. Washing in 75% occurred followed by drying in a speed vacuum and resuspension in DEPC-treated H<sub>2</sub>O.

## 2.11 Liquid Chromatography tandem Mass Spectrometry (LC-MS/MS)

### 2.11.1 Method development and optimisation

Pure nucleoside standards (C (Berry & Associates, PY 7215), A (Berry & Associates, PR 3005), G (Berry & Associates, PR 3703), U (Berry & Associates, PY 7780), 5mC (Berry & Associates,

PY 7637), 5hmC (Berry & Associates, PY 7596), 5fC (Berry & Associates, PY 7599), 5CaC (Berry & Associates, PY 7598) and m6A (Berry & Associates, PR 3732)) were resuspended in LC-MS grade H<sub>2</sub>O (Fisher, W/0112/17). These samples were initially tested at 500 fmol using a method adapted from Amouroux et al., 2016. Briefly, separation was performed on an Agilent RRHD Eclipse Plus C18 2.1 × 100 mm 1.8u column by using the UHPLC 1290 system (Agilent) and analysed using an Agilent 6490 triple quadrupole mass spectrometer. The chromatography programme is outlined in Table 2.4. This was adapted to a shorter program, also given in Table 2.4, to facilitate optimisation of sequential mass spectrometer parameters in series. In both cases buffer A consisted of 99.9% LC-MS/MS H<sub>2</sub>O (Fischer, W011217) and 0.1% formic acid (Sigma, 94318) and buffer consisted of 79.9% methanol (Fischer, 10653963), 20% LC-MS/MS H<sub>2</sub>O (Fischer, W011217) and 0.1% formic acid (Sigma, 94318). After optimisation was complete, the method was altered to the parameters outlined in Table 2.5, before limit of quantification and limit of detection values were determined using a standard curve with dilution ration 1:5:2. The standard curve used is provided in Table 2.6. LOQ and LOD thresholds were defined as signal to noise ratios of 10 and 3 respectively, as determined by the peak-to-peak method.

| Standard RNA protocol |       |        | RNA rapid screening protocol |       |       |
|-----------------------|-------|--------|------------------------------|-------|-------|
| Time [minutes]        | A [%] | B [%]  | Time [minutes]               | A [%] | B [%] |
| 0.00                  | 99.90 | 0.10   | 0.00                         | 97.00 | 3.00  |
| 4.00                  | 99.90 | 0.10   | 0.10                         | 97.00 | 3.00  |
| 14.00                 | 80.00 | 20.00  | 3.00                         | 40.00 | 60.00 |
| 14.50                 | 0.00  | 100.00 | 4.00                         | 97.00 | 3.00  |
| 15.50                 | 0.00  | 100.00 |                              |       |       |
| 16.00                 | 99.90 | 0.10   |                              |       |       |
| 20.00                 | 99.90 | 0.10   |                              |       |       |

Table 2.4 UHPLC mobile phase composition.

| Parameter                         | Before | Optimal |
|-----------------------------------|--------|---------|
| Sheath gas temperature (SGT) [°C] | 350    | 350     |
| Sheath gas flow (SGF) [l/min]     | 12     | 10      |
| Nozzle [V]                        | 0      | 0       |
| Nebulizer pressure [psi]          | 55     | 20      |
| Capillary [V]                     | 3000   | 1500    |
| High pressure RF (RFH) [V]        | 110    | 160     |
| Low pressure RF (RFL) [V]         | 60     | 80      |
| Gas temperature (DGT) [°C]        | 130    | 230     |
| Gas flow (DGF) [l/min]            | 20     | 11      |

Table 2.5 Mass spectrometry parameters.

| Levels | Dilution | C and A concentration [fmol/ $\mu$ l] | 5mC, 5hmC, 5fC and m6A concentration [fmol/ $\mu$ l] | C and A amount injected [fmol] | 5mC, 5hmC, 5fC and m6A amount injected [fmol] |
|--------|----------|---------------------------------------|------------------------------------------------------|--------------------------------|-----------------------------------------------|
| 10     | 1:20     | 500                                   | 25                                                   | 5000                           | 250                                           |
| 9      | 1:5      | 100                                   | 5                                                    | 1000                           | 50                                            |
| 8      | 1:2      | 50                                    | 2.5                                                  | 500                            | 25                                            |
| 7      | 1:5      | 10                                    | 0.5                                                  | 100                            | 5                                             |
| 6      | 1:2      | 5                                     | 0.250                                                | 50                             | 2.50                                          |
| 5      | 1:5      | 1                                     | 0.05                                                 | 10                             | 0.5                                           |
| 4      | 1:2      | 0.5                                   | 0.025                                                | 5                              | 0.25                                          |
| 3      | 1:5      | 0.1                                   | 0.005                                                | 1                              | 0.05                                          |
| 2      | 1:2      | 0.05                                  | 0.0025                                               | 0.5                            | 0.025                                         |
| 1      | 1:5      | 0.01                                  | 0.0005                                               | 0.1                            | 0.005                                         |

**Table 2.6 Nucleoside standard curve for LC-MS/MS analysis.**

After optimisation, heavy-labelled nucleoside isotopologues (C-13C915N3 (Cambridge Isotope Laboratories, CNLM-3807-PK), 5mC-13CD3 (Santa Cruz, SC-482515), 5hmC-13D2 (Toronto Research Chemicals, H947092), A-13C1015N5 (Cambridge Isotope Laboratories, CNLM-3806-CA) and m6A-D3 (Toronto Research Chemicals, M275897)) were resuspended in LC-MS grade H<sub>2</sub>O and used in addition to the light nucleoside standards to assess, and correct for, ion suppression.

## 2.11.2 Sample preparation

### 2.11.2.1 Enzymatic cleaning

A Microcon-10kDa Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, MRCPRT010) was used to clean an enzymatic mixture used to digest RNA samples to nucleosides (Nucleoside Digestion Mix (NEB, M0649S)) to reduce contamination, particularly

of glycerol and nucleosides. 75 µl 1x NEB buffer was applied to the column, which was spun at 14,000 xg for an hour at 8 °C in order to precondition it. 0.2 µl of the enzymatic digestion mix was taken per RNA sample to be digested, added to 50 µl 1x NEB buffer, and added to the column which was spun at 14,000 xg for 15 minutes at 8 °C. The enzyme was washed with a further 50 µl 1x NEB buffer and spun as before. Elution into a collection tube was performed by inverting the column and centrifuging at 3,000 xg for 1 minute at 8 °C. The elution volume was determined through the use of a top pan balance and used to calculate the appropriate volume of washed enzyme to use for each reaction. Once cleaned and eluted, the enzyme was used immediately for RNA digestion.

#### 2.11.2.2 RNA digestion to nucleosides

RNA samples up to 250 ng were digested for 12 hours at 37 °C in a reaction mix as specified in Table 2.7.

| Component                    | Volume [µl]  |
|------------------------------|--------------|
| RNA                          | x            |
| LC-MS/MS<br>H <sub>2</sub> O | 18 - (x + y) |
| 10x NEB buffer               | 2            |
| enzyme                       | y            |
| Total                        | 20           |

**Table 2.7 RNA digestion master mix for LC-MS/MS.**

#### 2.11.2.3 Sample preparation

The digested RNA nucleoside mix outlined above (section 2.11.2.20) was diluted 1:500 and 1:5 to enable the quantification of nucleosides over a very large abundance range. To all samples

and standard curve points heavy-labelled nucleoside isotopologues were added to a total amount equal to that of the light nucleosides in the level 6 standard curve point (50 fmol for unmodified nucleosides and 2.5 fmol for modified nucleosides). All samples and standards were run in technical duplicate.

## 2.12 Real-time quantitative PCR (RT-qPCR)

RT-qPCR was performed as described in Leitch et al., 2013. Briefly, reverse transcription was performed on 1 µg RNA using the Takara Clontech Primescript RT reagent kit with gDNA eraser (perfect real time) (Clontech, RR047A) with random primers. A fraction of all samples was taken to create a standard for downstream analysis before the remaining volume of samples was diluted 1:40. Of this dilution 3 µl was added to each 10 µl Bioline Sensimix SYBR No-Rox (Bioline, QT650-05) reaction mix and RT-qPCR performed using both the standard-curve and  $\Delta\Delta CT$  methods. Cycling conditions were 95 °C 10 min; 40 cycles of 95 °C 15 s; 60 °C 15 s; and 72 °C 15 s, followed by a melting curve from 55 °C to 95 °C. Primer sequences are provided in Table 2.8 along with references where applicable.

| Region            | Forward                 | Reverse                  | Annealing temperature [°C] | Reference           |
|-------------------|-------------------------|--------------------------|----------------------------|---------------------|
| TRDMT1            | GCAGAGGATGCGCAGATTGAG   | ATTGCCAAGCAGCCGGTAAC     | 60                         |                     |
| NOP2              | CCCATGCAGGAACTCGGCAG    | ACGGGCTTTCTTGGTGGGC      | 60                         |                     |
| NSUN2             | CAGCGACGGCGGCAGAAA      | CTTTGGCATGGCTTTTGTACCC   | 60                         |                     |
| NSUN3             | GCTGCAATGTGCTTATCCAGGT  | GGAGCGTCAACTAACACCTTGTC  | 60                         |                     |
| NSUN4             | AGAAATGGGCTACCACAGAACC  | GCTGCTCCAGCTTAGCACTCAC   | 60                         |                     |
| NSUN5             | CTGGGGTCTCTTGCTGCGAGT   | AGCATGCCAGAGCCACTACAAGA  | 60                         |                     |
| NSUN6             | CCCATGACTGGACCAAGGAAGAA | ATGGGCCCCCTCTTAACACTGCAT | 60                         |                     |
| NSUN7             | CTGTTCTCCCACTGTGCACCC   | CTCAGATGGGTCCCGCTCTCT    | 60                         |                     |
| GAPDH             | CTGACGTGCCGCCTGGAGAAA   | CCCGGCATCGAAGGTGGAAGA    | 60                         | Leitch et al., 2013 |
| DNMT1             | CGGAAGTAGGTGAAGTTTCAA   | CAAATAGATCCCCAAGATCCA    | 60                         |                     |
| DNMT3A1 & DNMT3A2 | TGATGATTGACGCCAAAGAA    | TCACAGTGGATGCCAAAGG      | 60                         |                     |
| DNMT3A1           | TTCTCCTCCTGTTCCTCTCCT   | GCCTACTGCCCAGCAATG       | 60                         |                     |
| DNMT3B            | GATTCTGGGGGAGGCTCTT     | CAGGACGCGCTCTGAAAG       | 60                         |                     |
| DNMT3L            | AACATTCCAGGGCTGAAGAG    | GCTTGCTCCTGCTTCTGACT     | 60                         |                     |
| TET1              | AGCACCCCGAGGACGATAAC    | CAAAGGAAGGATCGGCGTAG     | 60                         |                     |
| TET2              | TTTGTCAAACCTGCTGGCCAA   | GGGATGTCACTGCCTTCCTC     | 60                         |                     |
| TET3              | CTGAGGACGAGCAACTGCAC    | CGTTCTGGTTTTCTCTGCTG     | 60                         |                     |
| Rex1              | TCACTGTGCTGCCTCCAAGT    | GGGCACTGATCCGCAAAC       | 54                         | Kim et al., 2013    |
| Nr0b1             | TCCAGGCCATCAAGAGTTTC    | ATCTGCTGGGTTCTCCACTG     | 50                         | Kim et al., 2013    |
| Fgf5              | GCAGCCCACGGGTCAA        | CGGTTGCTCGGACTGCTT       | 53                         | Kim et al., 2013    |

|                               |                          |                         |    |                                                   |
|-------------------------------|--------------------------|-------------------------|----|---------------------------------------------------|
| Cdx2                          | AGTGAGCTGGCTGCCACACT     | GCTGCTGCTGCTTCTTCTTGA   | 53 | Chrysanthou et al., 2018                          |
| Elf5                          | ATTGCTCGCAAGGTTACTCC     | GGATGCCACAGTTCTTTCAGG   | 52 | Chrysanthou et al., 2018                          |
| Gata6                         | GAAGCGCGTGCCCTTCATC      | GTAGTGGTGTGGTGTGACAGTTG | 52 | Artus et al., 2011,<br>Alexandrovich et al. 2006  |
| Sox17                         | CTTCCCTACCAGGGACACGA     | ACTGCTTCTGGCCCTCAGGT    | 53 | Artus et al., 2011,<br>Vazquez-Zapien et al. 2016 |
| Zeb2                          | GAGGAAAGAGATGGCCACG      | GTCAGCAGTTGGGCAAAAG     | 50 | Chrysanthou et al., 2018                          |
| Snai2                         | CCTTTCTCTTGCCCTCACTG     | ACAGCAGCCAGACTCCTCAT    | 51 | Chrysanthou et al., 2018                          |
| Snai1                         | CTTGTGTCTGCACGACCTGT     | CTTCACATCCGAGTGGGTTT    | 50 | Chrysanthou et al., 2018                          |
| CDH2                          | GTCTGTGGAGGCTTCTGGT      | GTCTCGTCCACCTTGAAA      | 50 | Chrysanthou et al., 2018                          |
| Vim                           | ACTGCAGGAGCTGAATGACC     | CGCATCTCCTCCTCGTACAG    | 52 | Chrysanthou et al., 2018                          |
| Evx1                          | GCTGTGTATGCAGAGCAGGTAGA  | CCCCTAAATGGCTGGGATGT    | 52 | Wen et al., 2017                                  |
| Mus musculus brachyury<br>(T) | ACAGTCACAGACATTTCTGACAGC | TCACATAGATGGGGGTGACAC   | 51 | Wen et al., 2017                                  |
| Mesp1                         | ATCCCAGGAAAGGCAGGAA      | GAAGGTGCCAAGACCAAAGG    | 51 | Wen et al., 2017                                  |
| 18s rRNA                      | CTCAACACGGGAAACCTCAC     | CGCTCCACCAACTAAGAACG    | 51 |                                                   |
| 28s rRNA                      | AGGGATAACTGGCTTGTGGC     | GATGGTAGCTTCGCCCCATT    | 53 |                                                   |
| 5s rRNA                       | TCTCGTCTGATCTCGGAAGC     | AGCCTACAGCACCCGGTATT    | 51 | Uchida et al., 2010                               |
| GOF18ΔPE-EGFP                 | AACCACTACCTGAGCACCC      | ACCTCTACAAATGTGGTATG    | 55 | Yoshimizu et al., 1999                            |

**Table 2.8 Primers used for RT-qPCR.**

## 2.13 Bacterial work and *in vitro* transcription

Preparing agar plates, colony picking, and stab culture were performed as outlined on Addgene (<http://www.addgene.org>). Transformation of Dam/Dcm<sup>r</sup> competent *Escherichia coli* (NEB,C2925I) was performed as per supplier's protocol using a plasmid encoding for a Flag-tagged histone H2A in a pCRII-TOPO vector (Nashun et al., 2010).

### 2.13.1 Plasmid extraction

Plasmids were extracted using the QIAGEN QIAprep Spin Miniprep Kit (27106) as per protocol.

### 2.13.2 Plasmid sequencing

Plasmid sequencing was performed by Genewiz with primers to the SP6 (ATTTAGGTGACACTATAG) and T7 (TAATACGACTCACTATAGGG) promoters used.

### 2.13.3 *In vitro* transcription

*In vitro* transcription was performed using Ambion MEGAscript T7 Kit (AM1330), with sample purification carried out using the Ambion MEGAclean Kit (AM1908). Both kits were used as per protocol, with either 5mCTP (Trilink, N-1014) or 5hmCTP (Trilink, TD-FP01A-4) used in lieu of CTP to transcribe from the Flag-tagged histone H2A in a pCRII-TOPO vector plasmid (Nashun et al., 2010).

# Chapter 3: Development of methods and tools

## 3.1 Introduction

Historically advancements in the field of biological RNA modifications have been limited by the lack of sufficiently accurate and sensitive techniques. Techniques in analysing RNA modifications can typically be considered as those which map the location of modification sites within the transcriptome, and those which quantify the abundance of the modification. Site mapping necessitates sequencing with techniques, mostly adapted versions of established DNA sequencing, using either the susceptibility to chemical deamination or immunoprecipitation prior to sequencing. During bisulphite sequencing (Bis-seq) sodium bisulphite causes cytosines to be deaminated to thymines, with modified cytosines resisting deamination due to the additional chemical group (Huang et al., 2019). Antibodies raised against modified nucleotides can be used to pull down RNA with the epitope modification in RNA immunoprecipitation sequencing (RIP-seq) (Meyer et al., 2012). Identification of modified RNA sites generated by specific enzymes can be assayed by mutagenesis of the active site of RNA modification depositing enzymes to result in stable covalent intermediates upon catalysis. These structures can be pulled down using antibodies to the enzyme in methylated individual-nucleotide-resolution crosslinking and immunoprecipitation (miCLIP) (Hussain et al., 2013).

None of these techniques are truly quantitative in their assessment of RNA modification abundance. In order to analyse this critical aspect of modified RNA biology, dot blotting or ultimately liquid chromatography tandem mass spectrometry are required. These techniques, though iteratively improved over time, retain certain limitations which I shall aim to address before applying my optimised techniques to *in vitro* and *in vivo* reprogramming models.

### 3.1.1 Dot blotting

Dot blotting is an antibody-based molecular biology technique upon which a number of variations exist. It specifically refers to a hybridisation method to fix a desired chemical to a

membrane. In regards to nucleic acids, it involves a small volume of denatured sample being applied to a membrane and fixed in place. The membrane is then blocked, incubated with a primary antibody against the epitope of interest, before further washing and horseradish peroxidase (HRP)-conjugated secondary antibody incubation. After the final wash steps, enhanced chemiluminescence (ECL) reagent is used to generate a chemiluminescent signal which is detected to generate an image either on film or electronically. When dot blotting using polynucleotides, equal loading can be determined through methylene blue staining (Mishima et al., 2015, Hodge, 1998, Miao et al., 2016, Cui et al., 2017).

Linearity of signal, epitope specificity, and reproducibility of results are common issues encountered in dot blotting. Each shall be carefully considered and assessed during method development.

### 3.1.2 Liquid chromatography tandem mass spectrometry

#### 3.1.2.1 Principles of chromatography and Van Deemter equation

All forms of chromatography are based upon the separation of the components of a mixture based upon differing relative affinities for a mobile and stationary phase which vary in their physicochemical properties. Depending upon the system and application different chemistries are used for the stationary and mobile phases. For liquid chromatography (LC) or gas chromatography (GC) separation occurs inside a column which is packaged with beads coated in the stationary phase (Fig 3.1). Particles can take a range of different paths through the bead-packed column (as illustrated by the green, blue and red arrows in Fig 3.1), resulting in peak broadening. The efficiency of separation on a column can be modelled using the van Deemter equation:

$$H = A + B/\mu + C\mu$$

Where:

$H$  = Height equivalent theoretical plate (HETP)

$\mu$  = average linear velocity of the mobile phase

$A$  = eddy diffusion term

$B$  = Longitudinal diffusion term

$C$  = Resistance to mass transfer (non-equilibrium) term

### Equation 3.1 The van Deemter equation

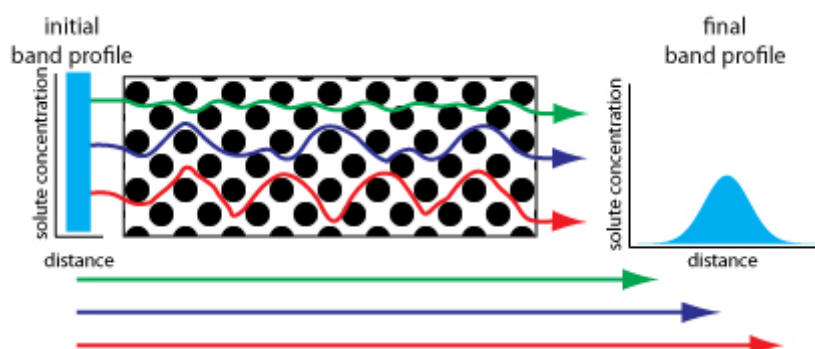


**Fig 3.1 Eddy diffusion in a packed chromatography column.**

Differing particle paths (green, blue or red) result in differing travel distances for elute particles. Adapted from Harvey, D, 2014. Figure reproduced under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) License.

HETP is a measure of the resolution of a peak for an eluted chemical. The total length of the column divided by this hypothetical measure of plate height gives the plate number ( $N$ ). This is effectively the maximum number of fully resolved peaks which can stretch from one end of the column to the other, and so  $N$  represents the separation efficiency of the column. Maximising

the N value is often the goal of chromatographic method development, and so for a given column length the lowest H value is optimal. Eddy diffusion (A term) occurs due to multiple flow paths through the packed column (Fig 3.1), while longitudinal diffusion (B term) is in the plane of the mobile phase flow (van Deemter et al., 1956). Resistance to mass transfer is a constant which factors in the non-uniformity of compound partition between the mobile and stationary phases (van Deemter et al., 1956). Factors which result in this non-uniformity include slower flow rates around the perimeter of the column and the kinetics of establishing partition equilibrium at the peak front (van Deemter et al., 1956). When generating any method involving liquid or gas chromatography these factors must be considered on an application-specific basis to determine and improve suitability. Due to RNA nucleosides being highly soluble, naturally in the aqueous state from their biological source, and being highly polar, reverse phase liquid chromatography is typically used to separate them. This involves a hydrophobic stationary phase (most commonly a C-18 column consisting of many 18-carbon long octadecylsilyl molecules covalently bonded to silica beads) over which an aqueous mobile phase is run. The mobile phase may be constant if the system is run in an equilibrium mode, however more commonly elution is achieved by the use of a gradient approach by which the mobile phase steadily becomes more hydrophobic.

### 3.1.2.2 Electrospray ionisation and triple quadrupole mass spectrometry

Mass spectrometry is an analytical process which uses the trajectory of charged molecules in an electric field to determine the mass to charge ratio ( $m/z$ ) of the molecule. Following separation of compounds within a mixture by high-performance liquid chromatography (HPLC), the eluate can run into a mass spectrometer for analysis. As ions are required for the detection by mass spectrometry, buffers are used to ensure that the analyte compounds of interest exist as either positive or negative ions. For the work outlined within this thesis positive mode mass

spectrometry was performed, with cations generated through the addition of 0.1 % formic acid to all mobile phase components (Chapter 2: Materials and Methods: Table 2.4). Desolvation of the charged aqueous analyte ions occurs in the nebuliser. Under pressure, and surrounded by a concentric flow of nitrogen gas, droplets of analyte-containing solvent undergo solvent evaporation, concentrating the charged molecules. This results in a coulombic explosion to give desolvated gas phase ions (Fenn et al., 1990, Mora et al., 2000). This process is referred to as electrospray ionisation (ESI) and is necessary to enable the movement, collisions, and eventually detection of the analyte compound by the downstream mass spectrometry instrumentation.

From the ESI source the charged ions pass through a hexabore capillary over which a constant potential difference is applied. This acts to filter out any neutral molecules before reaching the first  $m/z$  filter. There are a number of differing mass spectrometry instrumentations used to fulfil the role of a  $m/z$  filter. The importance of the strengths and limitations of each is determined by the desired application of the operator. Triple quadrupole mass spectrometry is outlined here, as it is the most common instrumentation chosen to analyse RNA modifications, representing the most sensitive instrumentation currently developed (Su et al., 2014, Huang et al., 2016, Yan et al., 2013, Huber et al., 2015, Fu et al., 2014). This is appropriate, as RNA modifications are often present at low abundances, particularly when compared to unmodified nucleosides, in biological samples (Huber et al., 2015). Sensitivity was considered to be the most important factor when optimisation was performed. The need for this was particularly acute, as it was hoped to be able to quantify the modified ribonucleoside levels within PGCs, which exist at a low cell number within the developing mouse embryo.

In a quadrupole both a direct current and a separate oscillating alternative current are applied between 4 parallel electrodes, which results in a particle of a specific  $m/z$  value passing through in a spiralling trajectory to exit to a detector (Fig 3.2) (Batey, 2014, Miller et al., 1986). Such an ion is referred to as a resonant ion. An ion of any other mass to charge ratio will result in a non-

resonant, unstable trajectory, which does not pass through the exit slit. By this methodology, only particles of a specified  $m/z$  value are able to pass through the quadrupole, thus acting as a  $m/z$  filter. Triple quadrupoles involve ions passing through 3 quadrupoles in tandem before reaching the detector. Quadrupole 1 specifies the precursor ion, while in quadrupole 2 collision induced dissociation (CID) occurs, whereby the ion fragments in a reproducible manner upon collision with inert gas molecules (Yost et al., 1979). The charged fragment(s) are then specified by the third quadrupole. By selecting the precursor ion, collision conditions, and product ion, the three quadrupoles act together to give a highly specific methodology for determining the identity of the original eluate from the HPLC column. By cycling between differing Q1 and Q3  $m/z$  values, multiple eluates can be sampled at once in a process referred to as multiple reaction monitoring (MRM). After quadrupole 3 the fragmented ion passes to a channel electron multiplier detector, which records an electrical impulse in proportion to the number of charged molecules which strike it (Yost et al., 1979).

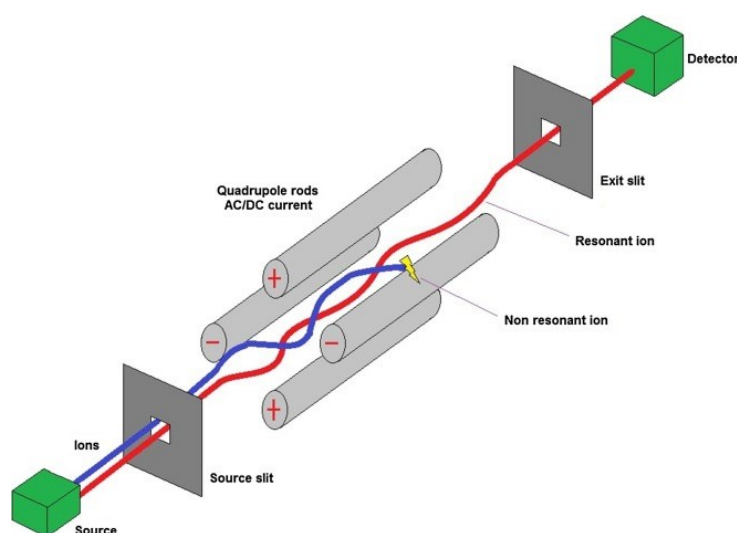


**Fig 3.2 Quadrupole mass spectrometer. Diagram of instrument operation.**

(Santoienma, 2018). Figure reproduced with permission of the rights holder, Elsevier.

## 3.2 Experimental aims

The overriding aim of this chapter was to develop two separate, but complementary, assays to quantify modified RNA nucleoside abundance in heterogeneous biological RNA samples. The diversity of RNA subtypes, along with material isolated from different *in vitro* and *in vivo* models, contribute to the heterogeneity of the samples. I resolved to optimise the sensitivity of RNA dot blotting, while ensuring that reproducible and substrate-specific results are obtained. Additionally, LC-MS/MS technology was to be used to build upon the intrinsic specificity of this assay to generate as sensitive an assay as could be achieved.

## 3.3 Results

### 3.3.1 Dot blotting development as a semi-quantitative measure of modified RNA nucleotide abundance

Optimisation began upon RNA extracted from J1 mESCs grown in serum containing media. This sample was chosen as a readily available source of RNA which is reflective of the type of sample the technique was to be used for after optimisation. The 33D3 anti-5mC antibody used as the primary antibody recognises both DNA 5mdC and RNA 5mC. As such, DNase treatment was performed on the samples by incubation with TURBO DNase for 20 minutes at 37 °C. After column clean-up, sample denaturation was compared between two established techniques. Denaturing either occurred by heating at 65 °C in 8M urea, or for 5 minutes in a formamide and formaldehyde solution containing 65.0% (v/v) formamide, 7.77 % formaldehyde (w/v) in 1.33x MOPS (hereby referred to as F&F) (Mishima et al., 2015, Hodge, 1998). As a positive control J1 mESC DNA was also used with denaturing performed through the addition of 0.1M NaOH and heating at 95 °C for 5 minutes. DNA 5mC dot blotting has been performed in the lab previously using this approach (Amouroux et al., 2016). The F&F methodology involves diluting the sample 8x overall though the use of F&F solution and 20x SSC, whereas the urea method

only results in a 2x dilution. The two methods were thus compared by both dotting the same amount of RNA on the membrane, and also dotting 4x more urea-treated sample. This enabled separate comparisons between equally loaded RNA, and the maximum amount of RNA that could be loaded from a set concentration of sample. The DNA positive control gives the strongest signal, even though the total amount of DNA is lower than that for the RNA. A greater signal is observed in the F&F treated sample than the urea treated sample at either 826.6 ng or 206.7 ng (Fig 3.3). Accordingly, the F&F method was taken forward for further work.



**Figure 3.3 Trialling RNA dot blot denaturation methodology shows F&F to be optimal.**

Two competing RNA denaturation methodologies were trialled together, alongside a DNA positive control, to determine which methodology to use for continuing optimisation. The data represents 1 biological replicate.

In order to test the linearity of the signal observed, a dilution series of RNA extracted from E14 mESCs and DNMT triple knockout (TKO) mESCs, each grown in either serum containing or 2i media, as well as EpiSCs were made. Methylene blue staining was used to control for any loading inconsistencies. Figure 3.4A and B show that there is a linear chemiluminescent 5mC signal detected for the E14 cell lines as well as the EpiSC across the concentration ranges tested.

No signal was observed in the TKO cell lines regardless of culturing media condition. Dot blotting was also performed using an anti-5hmC antibody (Fig. 3.4). Promisingly, a linear signal was also observed in this instance, although it was only in RNA from E14 serum cells that any signal was detected. Together, these results were highly promising in terms of the application of the technique, as well as the potential biological implications of these results. As I had observed a much stronger signal per nanogram of material for DNA than RNA with the anti-5mC antibody (Fig 3.1), I wanted to ensure that the signals observed were not the result of contaminating DNA. To ensure that the signal I was observing was not due to contaminating DNA 5mdC, I repeated the experiment with altered DNase treatments. Figure 3.5A shows clearly that the signal found in my initial samples was not detected upon either a new lot of TURBO DNase or QIAGEN RNase-free DNase digestion. This is consistent with the RNA 5mC signals detected so far being due to contaminating DNA 5mdC. Consistent with this, RNA samples treated with no nuclease give 5mC signal, which is lost upon DNase treatment (Fig 3.5B and Fig.3.5C). Digestion with DNase and RNase A ablates the methylene blue signal, indicating that RNA is present on the membrane in the absence of RNase A treatment (Fig 3.5B and Fig.3.5C). The methylene blue signal between the untreated and DNase treated samples are equivalent, showing that the contaminating DNA 5mdC giving rise to the detected signal represents a minor fraction of the total polynucleotide. As a result, even small amounts of contaminating DNA can dominate any signal that may come from RNA. This problem is exacerbated two-fold; cytidine methylation at the 5' position is more abundant in DNA than RNA, and, after factoring this in the 33D3 anti-5mC antibody appears to show greater affinity towards the DNA epitope (Fig. 3.1). This was later confirmed once a parallel LC-MS/MS approach was operational. An additional complication is that seemingly upon complete DNA removal the assay, as it existed at this point in time, was unable to detect a signal from RNA (Fig 3.5). To circumvent this and enable assay development, *in vitro* transcription using 5-methylcytidine triphosphate (5mCTP) or 5-

hydroxymethylcytidine triphosphate (5hmCTP) in lieu of CTP was performed to generate an artificial RNA transcript in which 100% of all cytidines are 5mC or 5hmC respectively. From this it was hypothesised that a signal could be detected, and optimisation performed, before reverting back to biological RNA samples. As an additional control, the RNA generated would be split into two equal volumes treated either with or without DNase-free RNase A. Any signal observed in the -RNase A sample, but either devoid or diminished in the +RNase A sample represents bona fide RNA 5mC or 5hmC as appropriate.

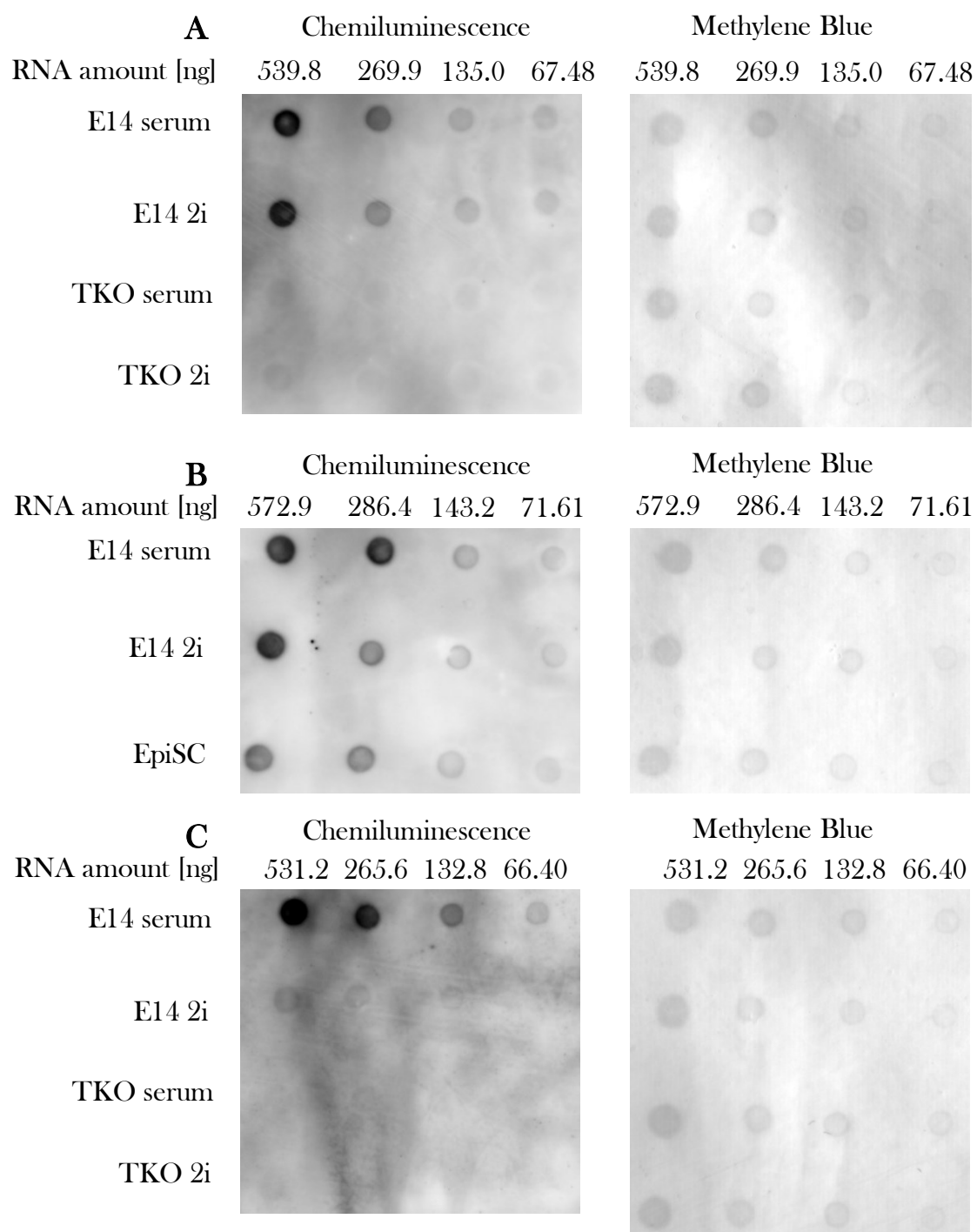


**Fig 3.4 Validation of signal linearity in RNA 5mC and 5hmC dot blotting.**

Anti-5mC (**A**) and (**B**) or anti-5hmC (**C**) dot blotting was performed on serial dilutions of RNA from a variety of mouse stem cells. The data represents 1 biological replicate.

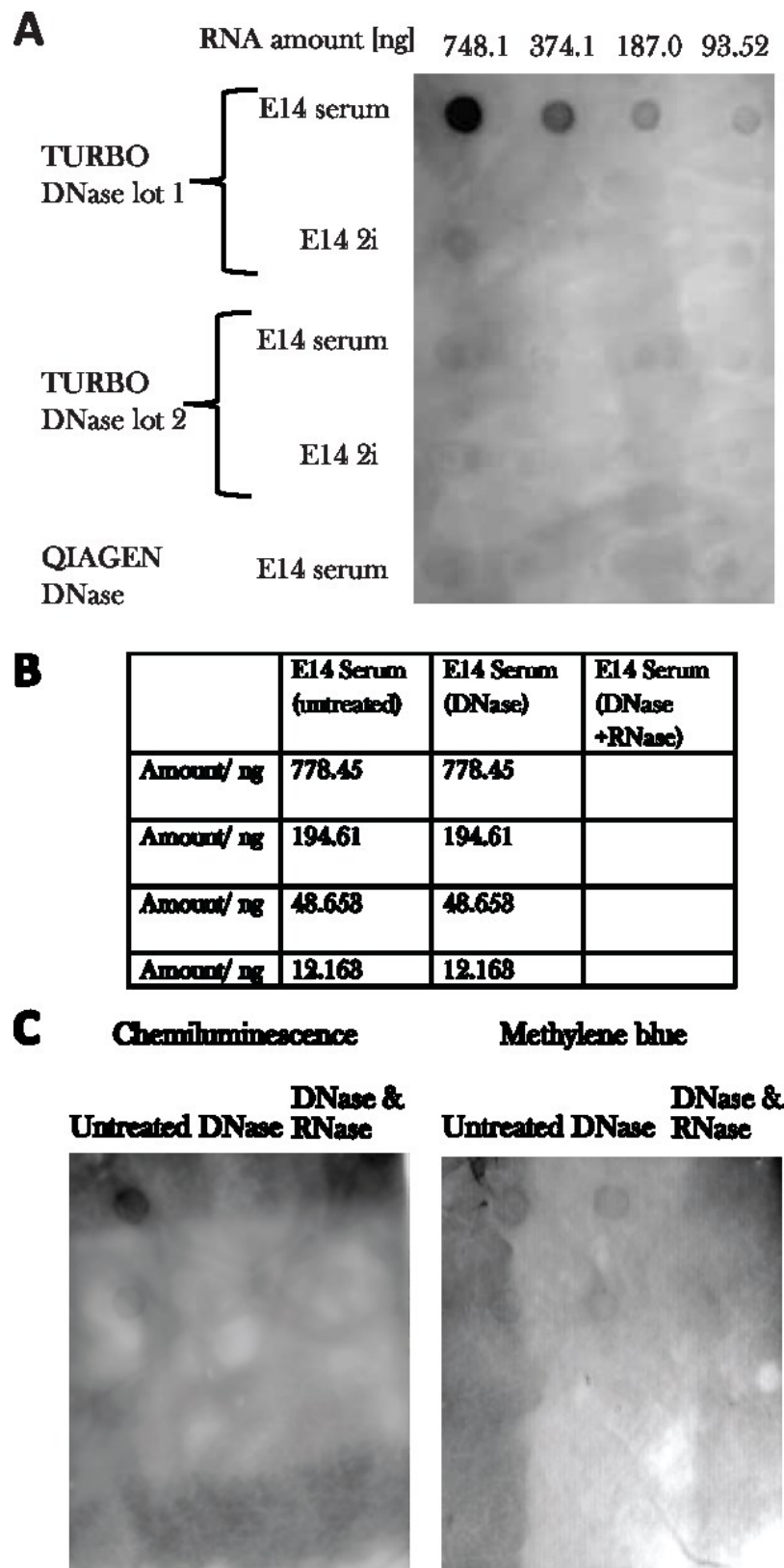


**Fig 3.5 RNA 5mC dot blot signal is due to contaminating DNA 5mdC.**

(A) Novel RNA preparations with fresh TURBO DNase or the alternative QIAGEN DNase result in no detectable RNA 5mC signal. The samples used and amounts of RNA blotted for the dot blot (C) are provided in table (B). In fresh preparations only the absence of DNase treatment causes 5mC signal to be detected. The data represents 1 biological replicate.

While generating *in vitro* transcribed RNA with 100% cytidine methylation or hydroxylation and performing digestion with/without RNase A, a divergent RNA dot blotting protocol was published which appeared promising (Miao et al., 2016). The key differences from the approaches I had been using were in denaturation and membrane fixing. RNA denaturation occurred with the addition of an equal volume of 2 M NaOH and incubation for 30 minutes at 4 °C and membrane fixing occurred by air drying for 15 minutes before baking at 80 °C for 30 minutes. This protocol was optimised using 100 ng of RNA extracted from mouse brain tissue. Figure 3.6 shows that for both 5mC (Fig. 3.6A) and 5hmC (Fig 3.6C) there is a signal detected which is ablated upon RNase A treatment. Despite background amounts of nucleic acids in samples treated with both DNase and RNase A were detected using a nanodrop, the methylene blue staining in the +RNase A samples appears identical to the -RNase A samples. Additionally, the signal detected was not a linear for either 5mC or 5hmC. Through repeated blotting and optimisation linear signals could not be achieved with this protocol. In light of this I decided at this stage to return to the F&F protocol. So far this method had not been proved to be able to detect RNA 5mC, although from DNA contamination it had been shown to be capable of providing the signal linearity necessary to enable semi-quantitation.

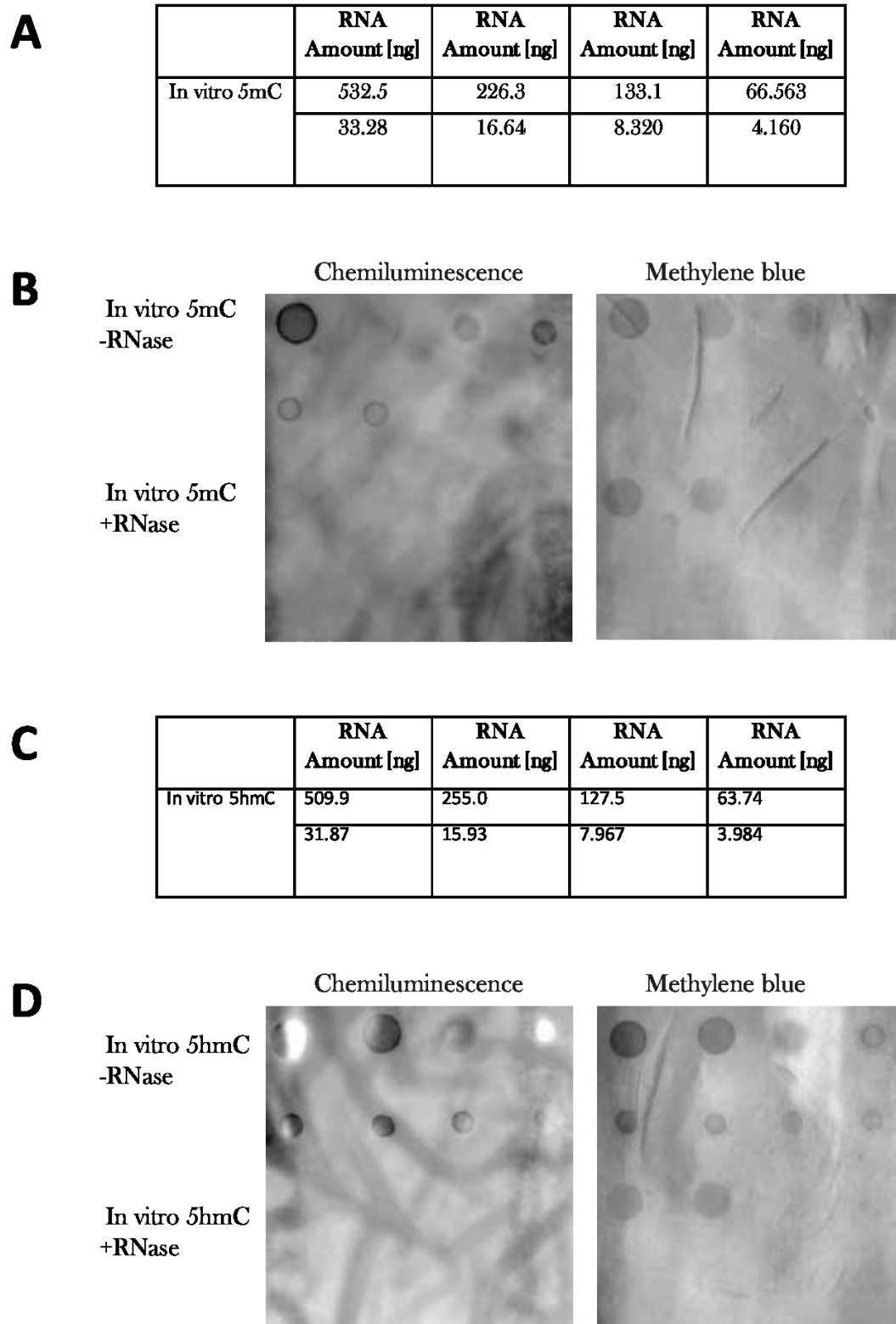


**Fig 3.6** *In vitro* transcribed 5mC and 5hmC dot blotting with RNase A control, NaOH denaturation, and baking.

Dot blotting was performed in accordance with Miao et al., 2016. The amount of RNA used in the 8-point standard curve for the 5mC (**B**) and 5hmC dot blots (**D**) are provided in tables (**A**) and (**C**) respectively. The data represents 1 biological replicate.

Upon returning to the F&F protocol, testing the *in vitro* transcribed modified transcripts was combined with an additional antibody dilution series for the 33D3 5mC. The optimal dilution factor was confirmed to be 1:1000, which was the dilution used up to this point based upon initial provisional trials with DNA contaminated samples (Fig. 3.4). This was the only dilution that gave a clearly definable circular dot signal over the background. A signal was detected, but only at around 1000 ng of total RNA. Given that biological samples will have RNA methylation levels approximately 100-fold lower than in these completely methylated transcripts, the amount of RNA required to generate a bona fide signal was unfeasibly high (Huber et al., 2015). The maximum amount of RNA that can be applied is defined by its solubility constant. At this stage it cannot be excluded that a trace of DNA 5mdC is contaminating the dot blots from the upstream RNA preparation. In order to conclusively rule this out, Dam-/Dcm- competent *Escherichia coli* were transformed with the Flag-tagged histone H2A in a pCRII-TOPO plasmid used for *in vitro* transcription (Nashun et al., 2010). These bacteria lack DNA methyltransferases Dam and Dcm, which result in them being unable to add or maintain DNA 5mdC (Marinus and Morris, 1973, and Geier and Modrich, 1979, May and Hattman 1975). After transfection puromycin selection was performed and colonies grown on agar dishes. The plasmids extracted from these bacteria had no DNA 5mdC, and so upon *in vitro* transcription with 5mCTP or 5hmCTP there was no capacity for contaminating DNA signal in the proceeding RNA dot blots. Figure 3.7 gives the first conclusively RNA-derived 5mC dot blotting signal. This signal is linear and absent upon RNase A treatment. It is slightly lower than that from *in vitro* transcribed RNA from the original plasmid which contained DNA 5mdC, showing that some DNA contamination was still occurring. The limit of detect was 133 ng of RNA, which was an improvement upon the work in figure 3.7. This was achieved through altering to blocking using milk as opposed to BSA. However, this level of sensitivity was not sufficient to be applicable to biological RNA samples.

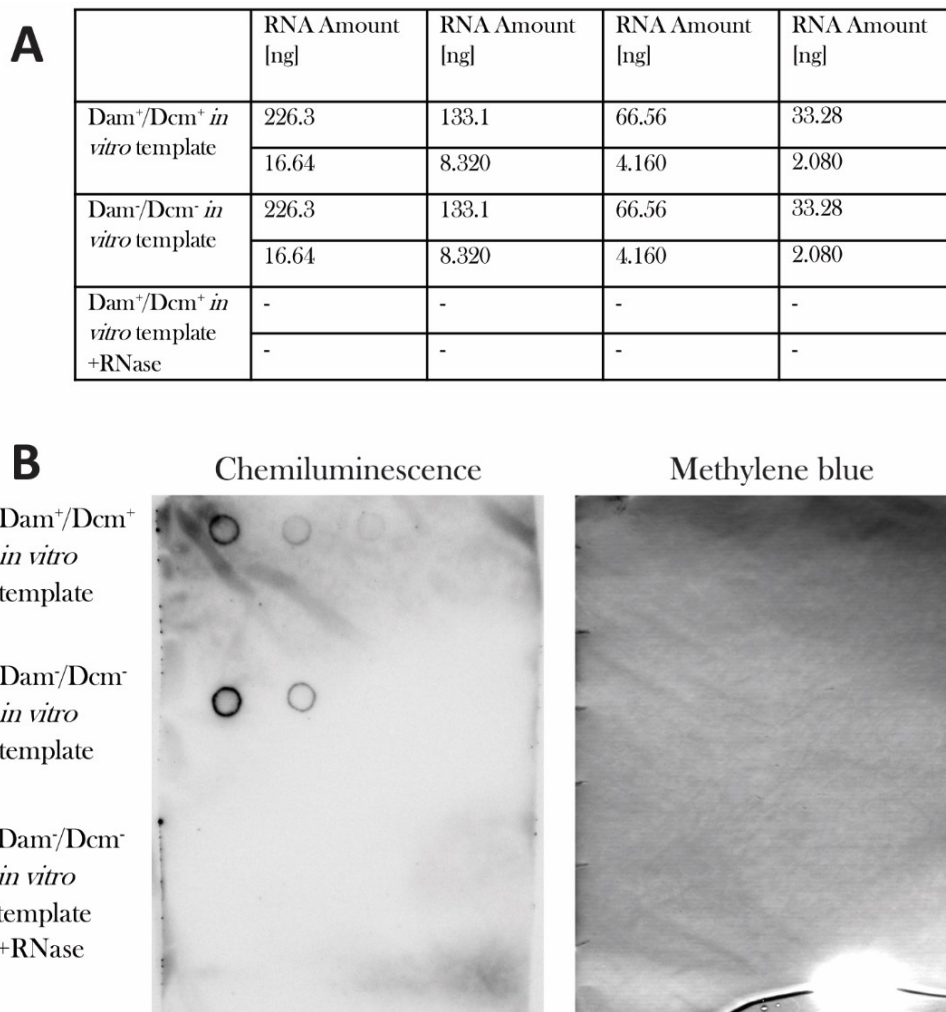


**Fig 3.7** *In vitro* transcribed RNA from an unmethylated plasmid gives a true RNA 5mC signal.

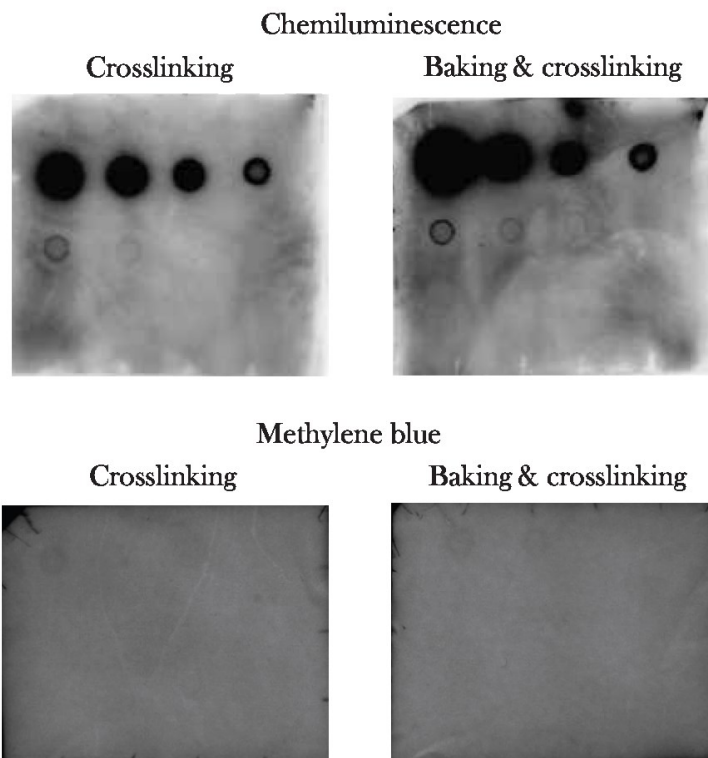
The amount of RNA used in the 8-point standard curve for the dot blot **(B)** is provided in table **(A)**. The data represents 1 biological replicate.

Fixation of RNA to membranes was found to be maximal upon baking at 80 °C for 30 minutes before UV crosslinking. This provided signal detection from 17ng of 100% methylated *in vitro* transcribed RNA when the signal from dots with a greater amount of RNA was allowed to oversaturate (Figure 3.8A and B). It follows that if RNA methylation exists at approximately 1 % in mouse tissues, that 1700 ng of mouse RNA would be necessary to detect a signal, which is eminently feasible (Huber et al., 2015). Figures 3.8A and B show that by employing the same approach for 100% hydroxymethylated *in vitro* transcribed RNA a bona fide RNA 5hmC signal can be detected from as little as ~0.12 ng of material. It was anticipated that RNA 5hmC would

exist at levels around 0.001% of all cytidines in mice, indicating that 12,000 ng of biological material would be required (Huber et al., 2015). It was unclear at this stage whether the RNA concentrations necessary to load this amount of RNA were physically possible. A strong brightfield signal is observed for the methylene blue stained image at RNA amounts that do not normally give a signal. This is the result of the intense chemiluminescent signals burning a permanent stain onto the membrane, as opposed to a methylene blue signal.

**A**

|          | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] |
|----------|-----------------------|-----------------------|-----------------------|-----------------------|
| Dam/Dcm  | 554.45                | 277.23                | 138.61                | 69.306                |
| in vitro | 34.653                | 17.327                | 8.6633                | 4.3316                |
| template | 2.1658                | 1.0829                | 0.54146               | 0.27073               |
| 5mC      |                       |                       |                       |                       |

**B****C**

|           | Amount<br>[ng] | Amount<br>[ng] | Amount<br>[ng] | Amount<br>[ng] |
|-----------|----------------|----------------|----------------|----------------|
| Dam/Dcm   | 255.0          | 127.5          | 63.74          | 31.87          |
| in vitro  | 15.93          | 7.967          | 3.984          | 1.992          |
| template: | 0.9959         | 0.4979         | 0.2490         | 0.1245         |
| 5hmC      |                |                |                |                |

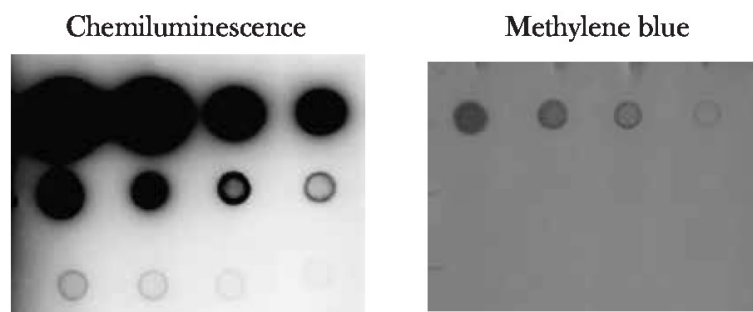
**D**

**Fig 3.8** Optimisation of RNA crosslinking and definition of  $\alpha$ -5mC and  $\alpha$ -5hmC RNA dot blot LOD values.

The amount of RNA used in the 12-point standard curve for the 5mC (**B**) and 5hmC dot blots (**D**) are provided in tables (**A**) and (**C**) respectively. The data represents 1 biological replicate.

Prior to returning to biological RNA transcripts, anti-5mC antibodies were trialled. The Millipore 33D3 clone (Millipore, MABE146) used to date is the most pervasively used in the field, however independently generated anti-5mC antibodies by Cell Signalling Technology (D3S2Z) and Abcam (RM231) were available at the time of investigation. These antibodies were compared to the 33D3 clone to determine if they were better suited to the application, particularly in light of the 33D3 clone showing epitope preference for 5mdC over 5mC (Fig. 3.1). The results of these trials are highlighted in Figure 3.9. The sensitivity of the Cell Signalling Technology antibody was lower than the 33D3 clone (Figure 3.9A and B). The Abcam RM231 antibody had a limit of detection similar to that of the Millipore 33D3 clone, however it did not display a linear signal (Figure 3.9C and D). The brightfield signal observed is primarily due to chemiluminescent burning as opposed to methylene blue signalling (Figure 3.9D). As these were the only antibodies commercially available, the 33D3 clone was advanced forward to use on biological RNA samples.

**A**

|                      | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] | RNA<br>Amount<br>[ng] |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Dam/Dcm              | 20                    | 10                    | 5                     | 2.5                   |
| in vitro<br>template | 1.25                  | 0.625                 | 0.313                 | 0.156                 |

**B**

Chemiluminescence

Millipore



Cell Signalling Technology



Methylene blue

Millipore



Cell Signalling Technology

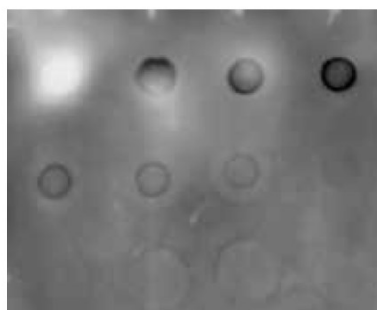
**C**

|           | Amount<br>[ng] | Amount<br>[ng] | Amount<br>[ng] | Amount<br>[ng] |
|-----------|----------------|----------------|----------------|----------------|
| Dam/Dcm   | 554.45         | 277.23         | 138.61         | 69.306         |
| in vitro  | 34.653         | 17.327         | 8.6633         | 4.3316         |
| template: | 2.1658         | 1.0829         | 0.54146        | 0.27073        |

**D**

Abcam

Chemiluminescence



Methylene blue

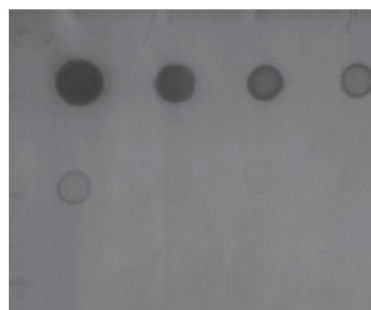


**Fig 3.9 Anti-5mC antibody trials reveal Abcam (RM231) to be the most sensitive antibody.**

The amount of RNA used in the 8-point standard curve for the dot blots in **(B)** is provided in table **(A)**. The amount of RNA used in the 12-point standard curve for the dot blot in **(D)** is provided in table **(C)**. The data represents 1 biological replicate.

Speed vacuuming was used to the point that RNA visibly precipitated from samples. Through this it was determined that 16,000 ng/μl was the upper RNA concentration threshold. This limited the amount of RNA that could be loaded in a single dot to 8000 ng. RNA from E14 and TKO mESCs grown in serum and 2i were used at this amount. However, no clear 5mC or 5hmC signal was detected from these samples (Figure 3.10). Only with the Abcam anti-5mC antibody could a faint signal be observed, although it was at background levels. *In vitro* transcribed RNA was used as a positive control in each case to show that the optimised methodology was still working. This demonstrates that the epitope abundances in these RNA samples was not sufficient to detect an RNA 5mC or 5hmC signal using this dot blot apparatus. This approach thus appeared to be at an impasse, until a novel anti-5mC antibody came onto the market. The Active Motif Abflex (91187) antibody was found to have a limit of detection in the region of 17-34 ng of 100% methylated *in vitro* transcribed RNA and displayed signal linearity (Figure 3.11). Once used with mESC RNA samples a discernible α-5mC signal could be observed using the Abflex antibody. The binding affinity and/or kinetics for this antibody are seemingly stronger than the other antibodies used. This exciting breakthrough facilitated the generation of the dot blot displayed in Figure 3.12. This shows RNA 5mC signal from RNA originating from differing mESC cell sates, which is linear and ablated upon RNase A treatment. As such, this methodology was now fit for purpose.

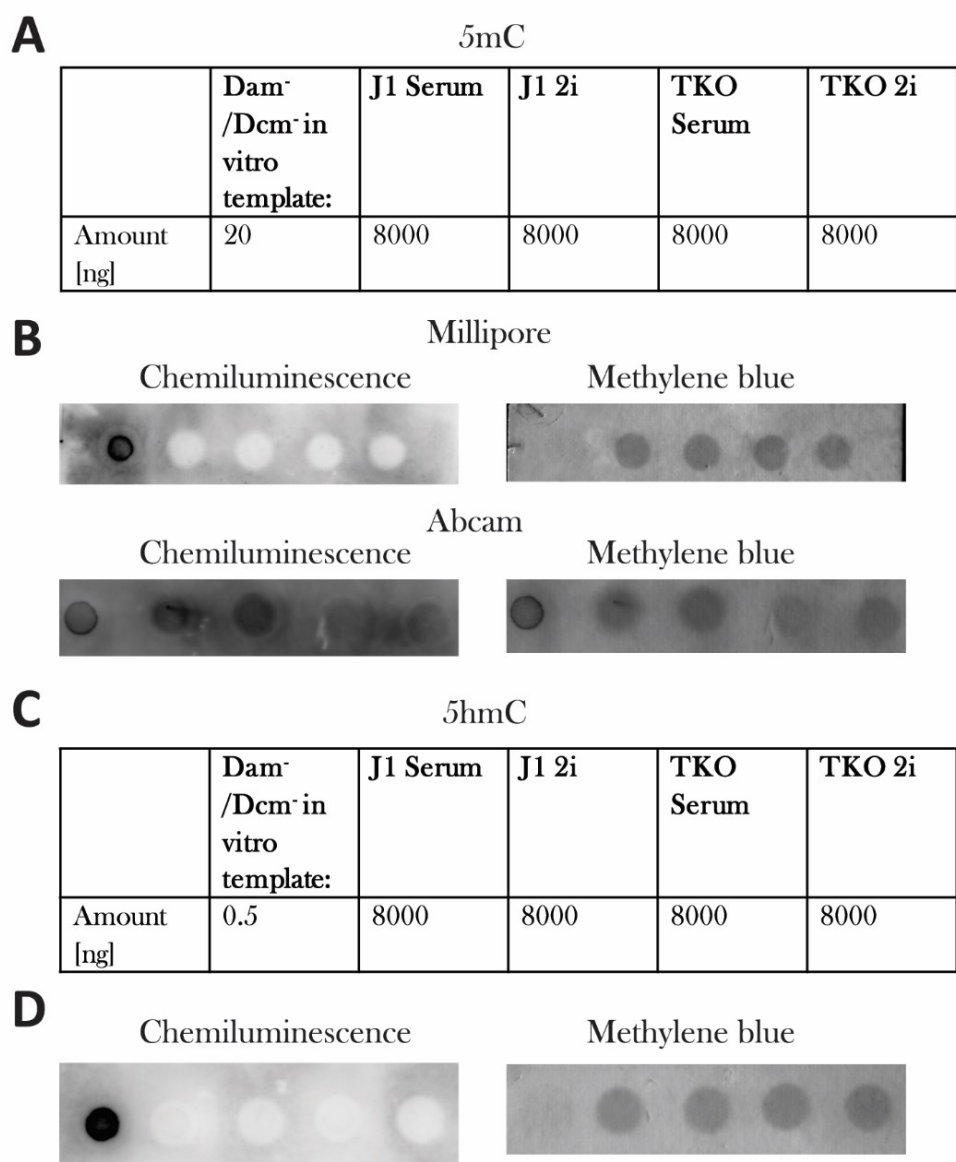


**Fig 3.10 5mC and 5hmC are undetectable in mESCs.**

The amount of RNA used in dot blots **(B)** and **(D)** are provided in tables **(A)** and **(C)**. The data represents 1 biological replicate.

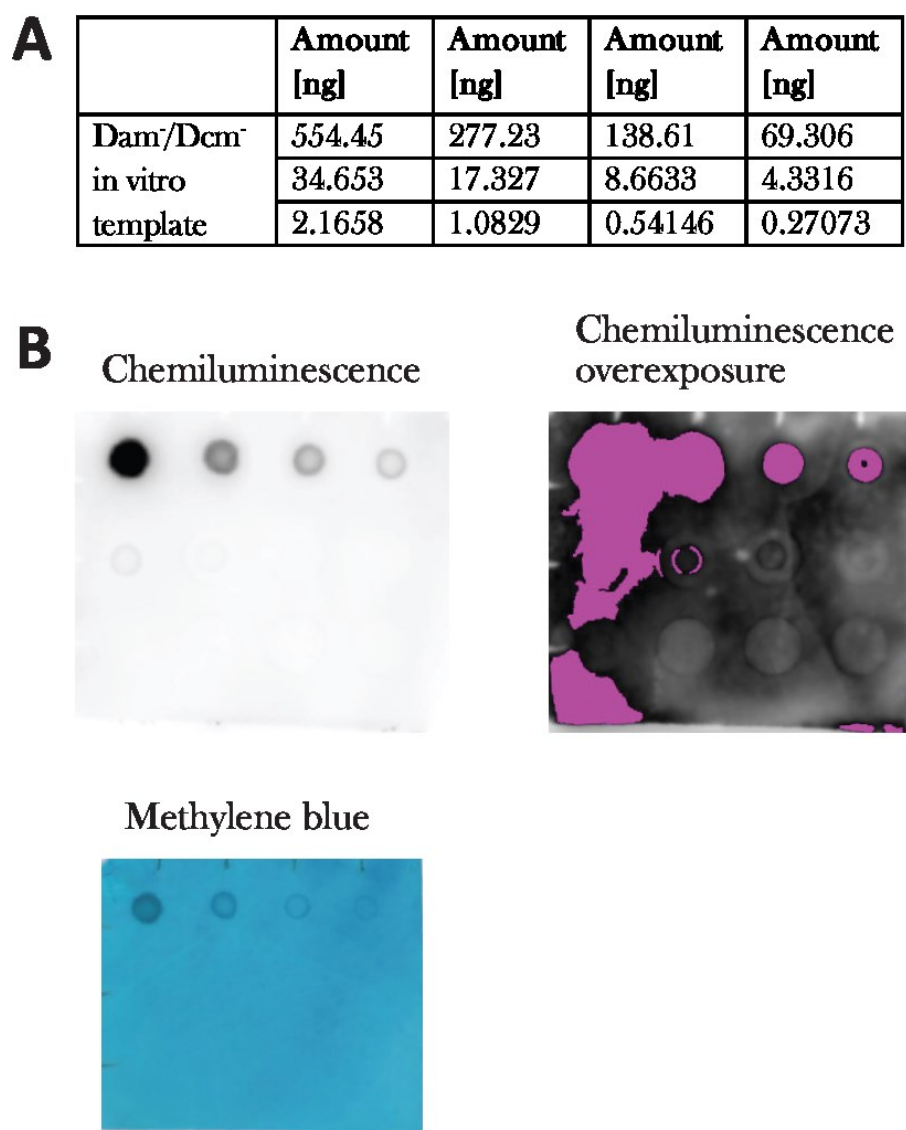


**Fig 3.11 Active Motif Abflex  $\alpha$ -5mC antibody trial and determination of LOD.**

The amount of RNA used in the 12-point standard curve for the dot blot in **(B)** is provided in table **(A)**. Areas of purple represent saturated pixels on the ImageQuant. The data represents 1 biological replicate.

**A**

| Cell line and media | Treatment     | Amount [ng] | Amount [ng] | Amount [ng] | Amount [ng] |
|---------------------|---------------|-------------|-------------|-------------|-------------|
| J1 Serum            | Untreated     | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase         | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase+ RNase  | -           | -           | -           | -           |
| J1 2i               | Untreated     | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase         | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase+ RNase  | -           | -           | -           | -           |
| TKO Serum           | Untreated     | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase         | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase + RNase | -           | -           | -           | -           |
| TKO 2i              | Untreated     | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase         | 337.4       | 84.35       | 21.09       | 5.272       |
|                     | DNase+ RNase  | -           | -           | -           | -           |

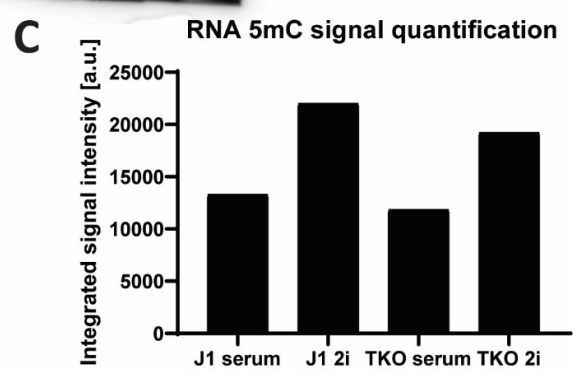
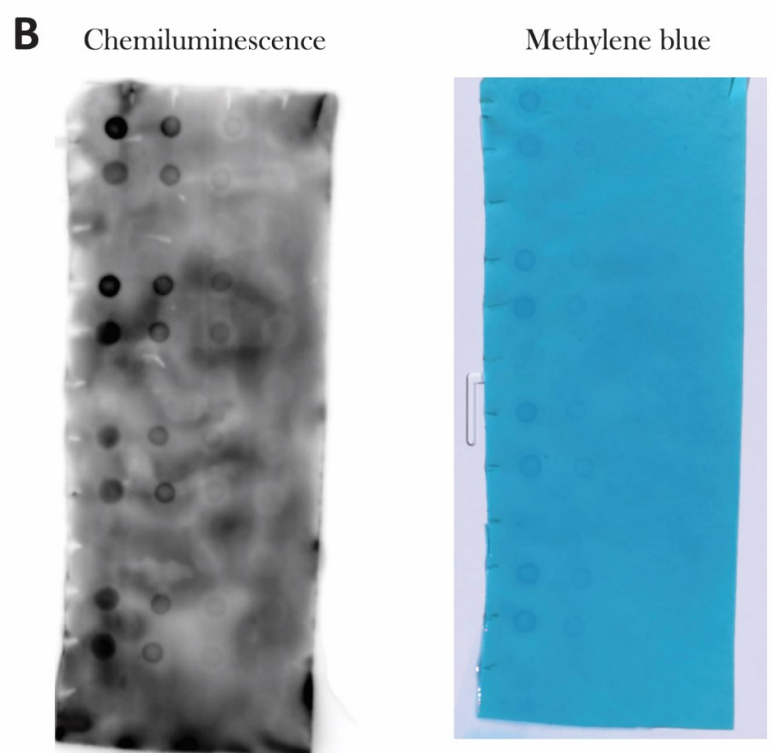


**Fig 3.12** Functionally optimised  $\alpha$ -5mC RNA dot blot gives a greater signal in ESCs grown in 2i than serum media.

The amount of RNA used in the 4-point standard curve for each of the samples used in the dot blots in **(B)** is provided in table **(A)**. **(B)** A representative dot blot is shown for 3 biological replicates. **(C)** Integrating the signal intensity over the dots enabled quantification of the 5mC signal.

### 3.3.2 LC-MS/MS optimisation quantitative measure of modified RNA nucleotide abundance

#### 3.3.2.1 LC validation

In parallel to dot blotting, a LC-MS/MS approach was developed and optimised to quantify modified nucleoside amounts. Chemically synthesised cytidine, guanosine, adenosine, uridine, 5-methylcytidine, 5-hydroxymethylcytidine, 5-formylcytidine and 5-carboxycytidine standards were resuspended in LC-MS/MS H<sub>2</sub>O. These standards represent the modified cytidine molecules of experimental interest in addition to the most abundant RNA nucleosides (unmodified cytidine, guanosine, adenosine and uridine). They were initially tested at 500 fmol using a method adapted from Amouroux et al., 2016 (Table 2.4). This amount of material was chosen to represent a greater analyte abundance than would be expected to be observed *in vivo* for the modified RNAs, without overloading the column. This reverse phase method was initially designed for DNA nucleoside separation. DNA and RNA nucleosides with equivalent organic base structures differ chemically only in the 2' hydroxyl group present in RNA. As such, RNA nucleosides were expected to be separated using this programme and elute earlier than their DNA counterparts. The chromatography programme is provided in Table 2.4. Separation was performed using an Agilent RRHD Eclipse Plus C18 2.1 × 100 mm 1.8u column on the UHPLC 1290 system. This was followed by analysis using an Agilent 6490 triple quadrupole mass spectrometer. The m/z values for the precursor and product ions were predicted using the known compound Mr values and hypothesising that fragmentation primarily occurred at the N-glycosidic bonds linking the ribose to the base (Table 3.1). The system was run in positive mode, with cations formed through the addition of 0.1 % v/v formic acid to the running buffers. Even analytical grade H<sub>2</sub>O contains ionic impurities, mostly leached from glassware. As a result of this the sodiated, potassiated, and protonated forms of all ions were considered (Table 3.1).

| Nucleoside                     | Cationic form             | Precursor [m/z] | Product [m/z] | Absolute abundance [a.u.] | Percentage abundance [%] | Retention time [minutes] |
|--------------------------------|---------------------------|-----------------|---------------|---------------------------|--------------------------|--------------------------|
| Cytidine (C)                   | <b>CNa<sup>+</sup></b>    | 266.1           | 134.1         | 6579749                   | 0.994191                 | 1.092                    |
|                                | CK <sup>+</sup>           | 282.1           | 150.1         | 15157                     | 0.00229                  |                          |
|                                | CH <sup>+</sup>           | 244.1           | 112.1         | 23288                     | 0.003519                 |                          |
| Guanine (G)                    | <b>GNa<sup>+</sup></b>    | 306.2           | 174.2         | 6753894                   | 0.96544                  | 5.647                    |
|                                | GK <sup>+</sup>           | 322.2           | 190.2         | 207759                    | 0.029698                 |                          |
|                                | GH <sup>+</sup>           | 284.2           | 152.2         | 34014                     | 0.004862                 |                          |
| Adenosine (A)                  | ANa <sup>+</sup>          | 290.2           | 158.2         | 50998                     | 0.014579                 | 4.697                    |
|                                | AK <sup>+</sup>           | 306.2           | 174.2         | 959                       | 0.000274                 |                          |
|                                | <b>AH<sup>+</sup></b>     | 268.2           | 136.2         | 3446016                   | 0.985147                 |                          |
| Uridine (U)                    | <b>UNa<sup>+</sup></b>    | 267.2           | 135.2         | 29164                     | 100                      | 2.411                    |
|                                | UK <sup>+</sup>           | 283.2           | 151.2         | N/A                       | 0                        |                          |
|                                | UH <sup>+</sup>           | 245.2           | 113.2         | N/A                       | 0                        |                          |
| 5-methylcytidine (5mC)         | <b>5mCNa<sup>+</sup></b>  | 280.1           | 148.1         | 11221334                  | 0.988316                 | 2.081                    |
|                                | 5mCK <sup>+</sup>         | 296.1           | 164.1         | 7782                      | 0.000685                 |                          |
|                                | 5mCH <sup>+</sup>         | 258.1           | 126.1         | 124878                    | 0.010999                 |                          |
| 5-hydroxymethylcytidine (5hmC) | <b>5hmCNa<sup>+</sup></b> | 296.1           | 164.1         | 5729420                   | 0.990708                 | 1.126                    |
|                                | 5hmCK <sup>+</sup>        | 312.1           | 180.1         | 10743                     | 0.001858                 |                          |
|                                | 5hmCH <sup>+</sup>        | 274.1           | 142.1         | 42993                     | 0.007434                 |                          |
| 5-formylcytidine (5fC)         | <b>5fCNa<sup>+</sup></b>  | 294.1           | 162.1         | 5783631                   | 0.995135                 | 5.917                    |
|                                | 5fCK <sup>+</sup>         | 310.1           | 178.1         | 12694                     | 0.002184                 |                          |
|                                | 5fCH <sup>+</sup>         | 272.1           | 140.1         | 15580                     | 0.002681                 |                          |
| 5-carboxycytidine (5CaC)       | <b>5CaCNa<sup>+</sup></b> | 309.2           | 177.2         | 81808                     | 0.908271                 | 1.871                    |
|                                | 5CaCK <sup>+</sup>        | 325.2           | 193.2         | 8262                      | 0.091729                 |                          |
|                                | 5CaCH <sup>+</sup>        | 287.2           | 155.2         | N/A                       | 0                        |                          |

**Table 3.1 Nucleoside cation abundance, transitions, and retention times.**

N/A represents ions that could not be detected above background. Abundances were calculated by integrating the chromatogram peak and are given in arbitrary units. The most abundant cation for each nucleoside is emboldened.

As Figure 3.13 and Table 3.1 show, all the tested standards were detected in at least 1 cationic form. In each instance there is a single dominant cationic form, being sodiated for all tested modifications with the exception of adenosine, which predominantly adopts the protonated form. Through a combination of differences in retention time and  $m/z$ , each nucleoside can be clearly separated from each other. This is vital to enable valid quantification of modified RNA abundances from the kind of heterogeneous mix a biological sample would represent. It was deemed at this stage that the chromatography method was thus sufficient to continue to optimisation of the mass spectrometry component of the assay. Should the chromatography process not have been sufficient upon the latter use of biological RNA extracts, iterative optimisation was planned.

Given that an equal number of moles of each nucleoside were injected during chromatographic optimisation, the abundance values vary substantially. This is a reflection of the established method being adapted for this application being optimised for the detection primarily of DNA 5mdC and 5hmdC. The carboxylic acid group of 5CaC has a  $pK_a$  value resulting in this functional group being negatively charged in some molecules, even in the presence of 0.1 % formic acid (Table 2.4). As a result, even after accepting a proton (or in most cases a sodium) at the basic amine group, the most common species is a neutral zwitterion. As neutral molecules are not detected by mass spectrometry, this explains the low abundance values. It was decided that this stage that 5caC could not be measured in the same positive-mode assay as the other modified cytidine derivatives and so was not considered during further optimisation. Attention was next turned to optimising the mass spectrometry parameters to achieve a highly sensitive and specific method.

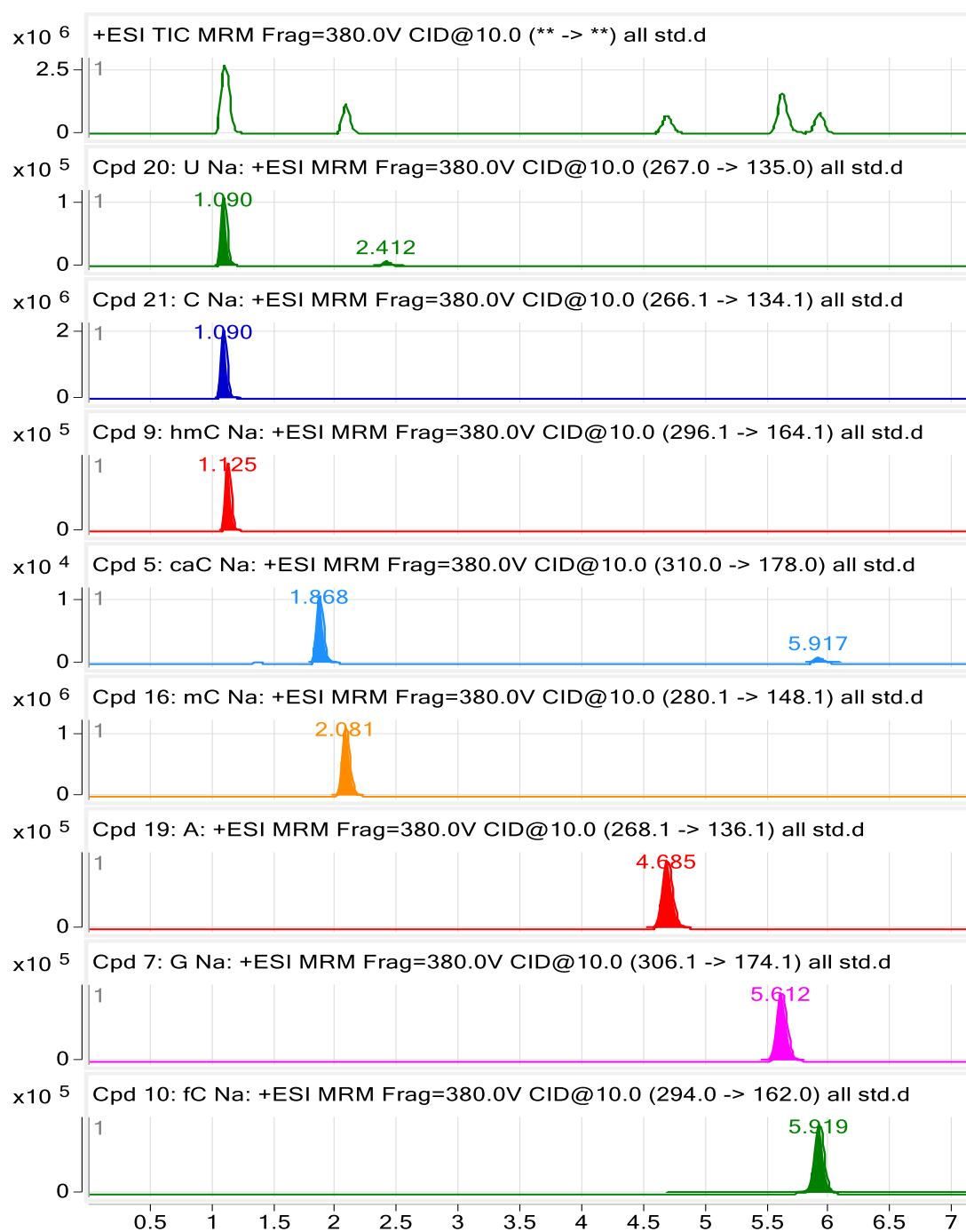


**Figure 3.13** Electrospray ionisation total internal chromatograms of nucleoside standards run together and in isolation.

Each chromatogram displays the most abundant cation for the respective nucleoside. The data presented is representative of  $n=3$  technical replicates from the resuspended nucleotides.

### 3.3.2.2 MS/MS method development

After showing that nucleoside standards could be separated and detected, attention turned to optimising the triple quadrupole mass spectrometry to give the maximum sensitivity. 5mC, 5hmC, 5fC and C were all considered independently. The critical adjustable parameters, and their starting values used thus far, are displayed in Table 3.2. Sheath gas temperature and flow relate to the inert nitrogen gas used to limit the spread of the eluate spray from the nebulizer during desolvation. A nebulizer pressure and nozzle voltage are applied at this point, before gaseous cations are extracted using a capillary voltage. Background contaminants are removed further using an iFunnel (Agilent). This piece of equipment involves two stages of electrical funnelling optics, firstly under a high pressure followed by low pressure, over which a voltage is applied. After precursor m/z selection in quadrupole 1, fragmentation of the ion occurs in quadrupole 2. This occurs due to the physical collision of the cation with gaseous particles. The gas temperature and flow rate into the quadrupole 2 can both be adjusted to give optimal fragmentation.

| Parameter                   | Initial value |
|-----------------------------|---------------|
| Sheath gas temperature [°C] | 350           |
| Sheath gas flow [l/min]     | 12            |
| Nozzle [V]                  | 0             |
| Nebulizer pressure [psi]    | 55            |
| Capillary [V]               | 3000          |
| iFunnel high pressure [V]   | 110           |
| iFunnel low pressure [V]    | 60            |
| Gas temperature [°C]        | 130           |
| Gas flow [l/min]            | 20            |

Table 3.2 The MS/MS parameters optimised with their initial values.

In order to decrease the time to optimise these parameters, the 20-minute chromatography run was shortened to a 4-minute rapid screening method given in Table 2.4. This cut the run time 5-fold, which was both advantageous in terms of preventing sample degradation as well as obtaining the necessary data in a timely fashion. Cytidine, 5-methylcytidine, 5-hydroxymethylcytidine and 5-formylcytidine standards were injected at 500 fmol for cytidine and 25 fmol for the modified nucleosides. 5mC and 5hmC were given primacy in terms of maximising the response and signal to noise, but it was also necessary to avoid selecting parameters which disproportionately reduce the sensitivity for the other cytidine nucleosides. In turn, each parameter was set to the lowest operating value and iteratively increased up until the instruments maximum. This was performed while holding all other parameters constant at their initial value. The voltages applied across the high pressure and low pressure funnels in the iFunnel function synergistically, and so were considered together (Figures 3.15-3.18). The data generated is provided in Figures 3.14-3.19. The data for cytidine is displayed separately from the modified nucleotides as its response values were much higher as a result of its greater abundance. Once this data was generated the optimal parameters for 5mC and 5hmC were determined (Table 3.3) and tested against the original method. Interestingly, the newly devised methods performed worse than the original method. By the nature of the software used it was not possible to change the parameters to the newly optimised parameters before moving onto the next parameter to be changed. When each alteration was added together, some aspect of their combination reduced the instrument sensitivity. This was hypothesised to be a result of sample degradation. In principal the spectrometry was optimised to give the greatest response for the amount of analyte present. However, through the degradation of the analyte the resulting reduced analyte abundances gave lower overall responses. Relative to the starting protocol optimised for DNA nucleosides, many of the parameters increased in severity, with higher temperatures, voltages and flow rates. These parameters were individually scaled down to give a

less damaging method that proved optimal for detection of both 5mC and 5hmC (Table 3.3). This method resulted in the limit of detection (LOD) and limit of quantification (LOQ) values, as defined by a signal to noise ratio of 3 or 10 respectively, outlined in Table 3.4. Demonstrable improvements upon the starting method for 5mC and 5hmC were achieved, which represents the most sensitive RNA modification detection technique developed to date. It is worth noting that the condition of the instrument when this experiment was performed was in the lower end of its normal operating range, and that in later experiments using biological material LOQ values were typically extended 5-fold lower than the values given in Table 3.4. When mESC RNA was used in accordance with this technique, quantification of each nucleoside was achievable (Figure 3.20). This showed the technique to be applicable to biological samples and not unduly affected by matrix effects resulting from the summation of contaminants remaining from the upstream sample preparation of cellular extracts.

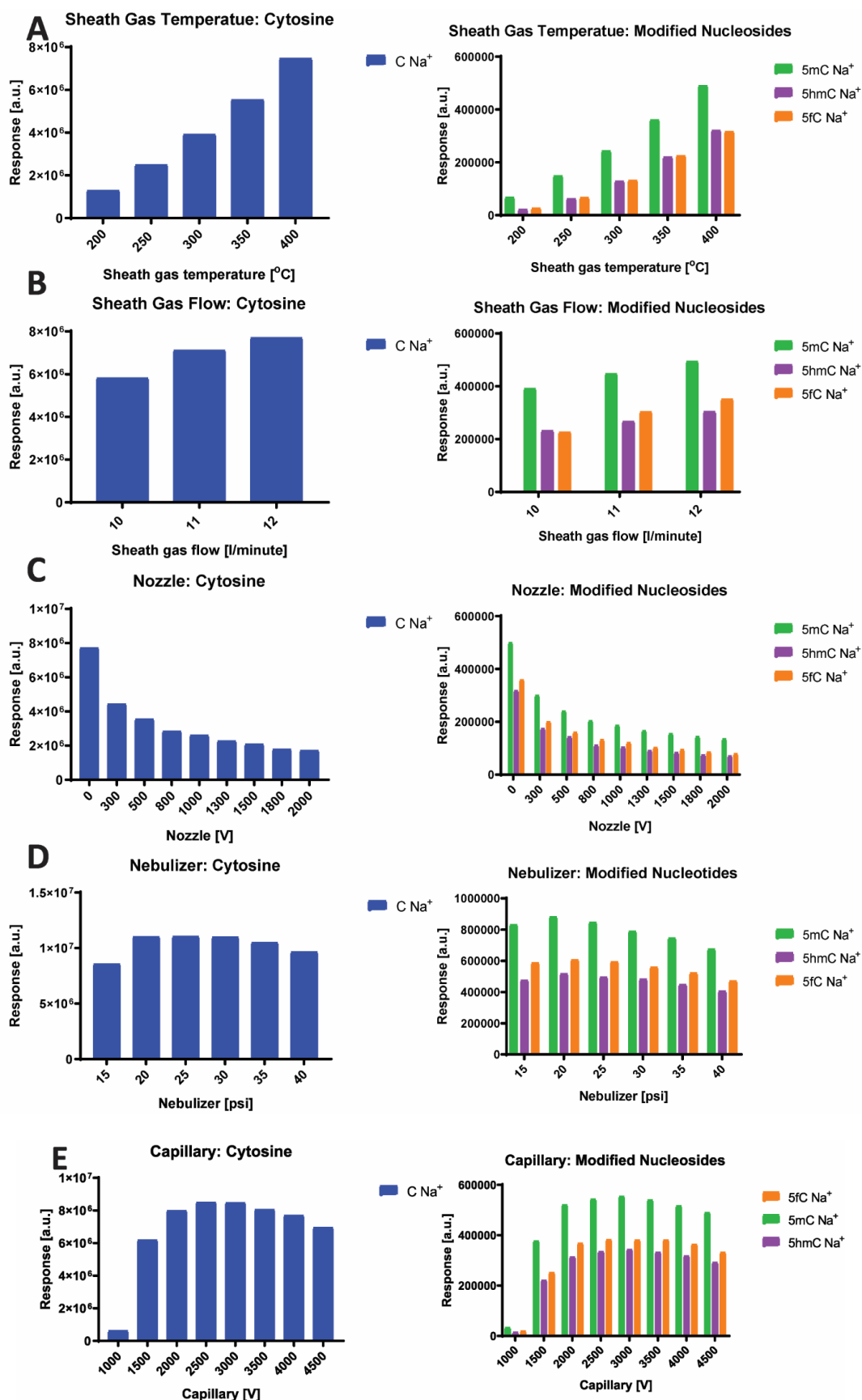


**Fig 3.14 Optimisation of mass spectrometry electrospray ionisation and cation extraction.**

The response values for assorted electrospray ionisation and cation extraction parameters are shown for cytidine and its modifications. The data presented is representative of n=2 technical replicates from the resuspended nucleotides.

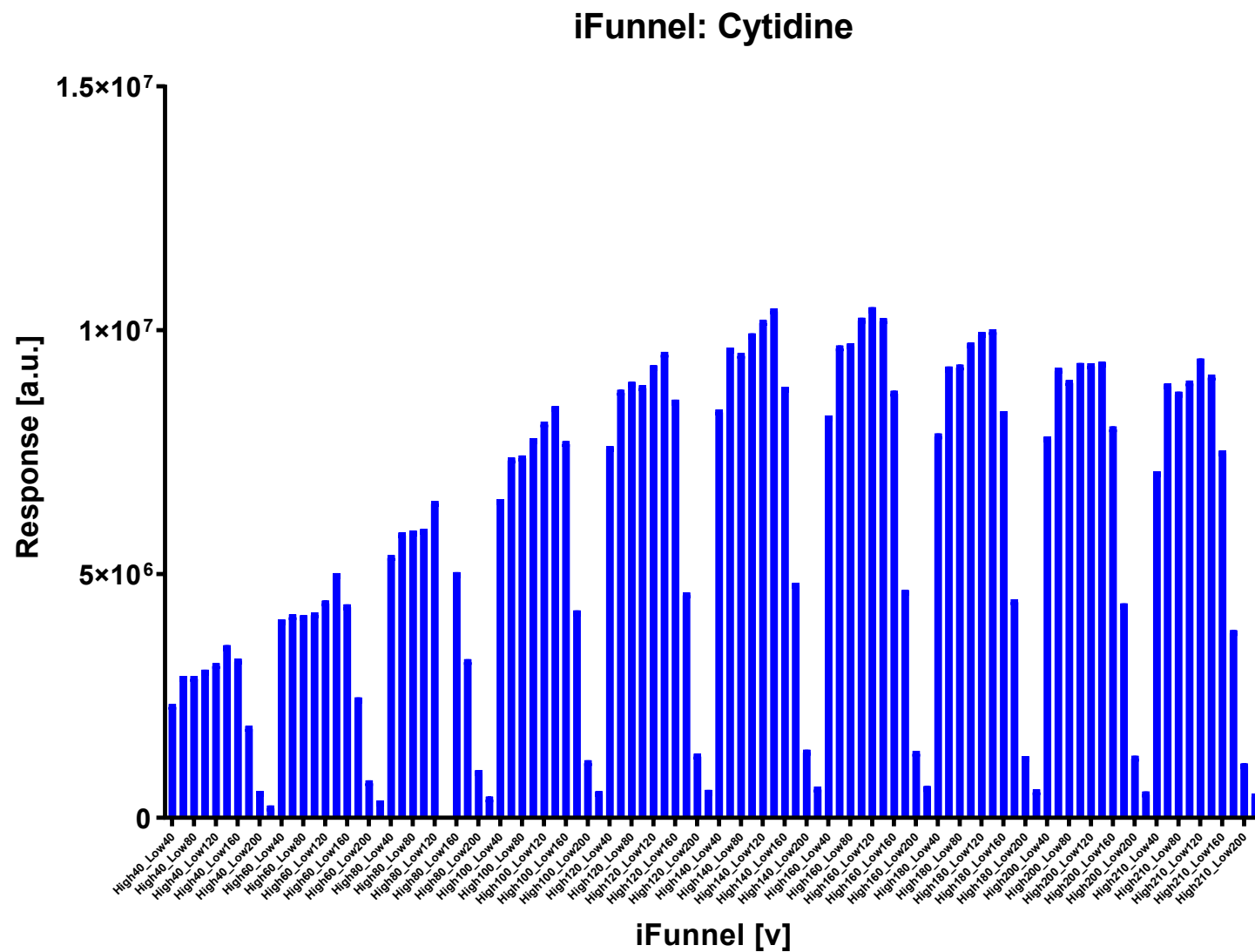


Fig 3.15 Optimisation of mass spectrometry iFunnel voltages for cytidine.

The response values for the iFunnel parameters are shown. The data presented is representative of n=2 technical replicates from the resuspended nucleotides.

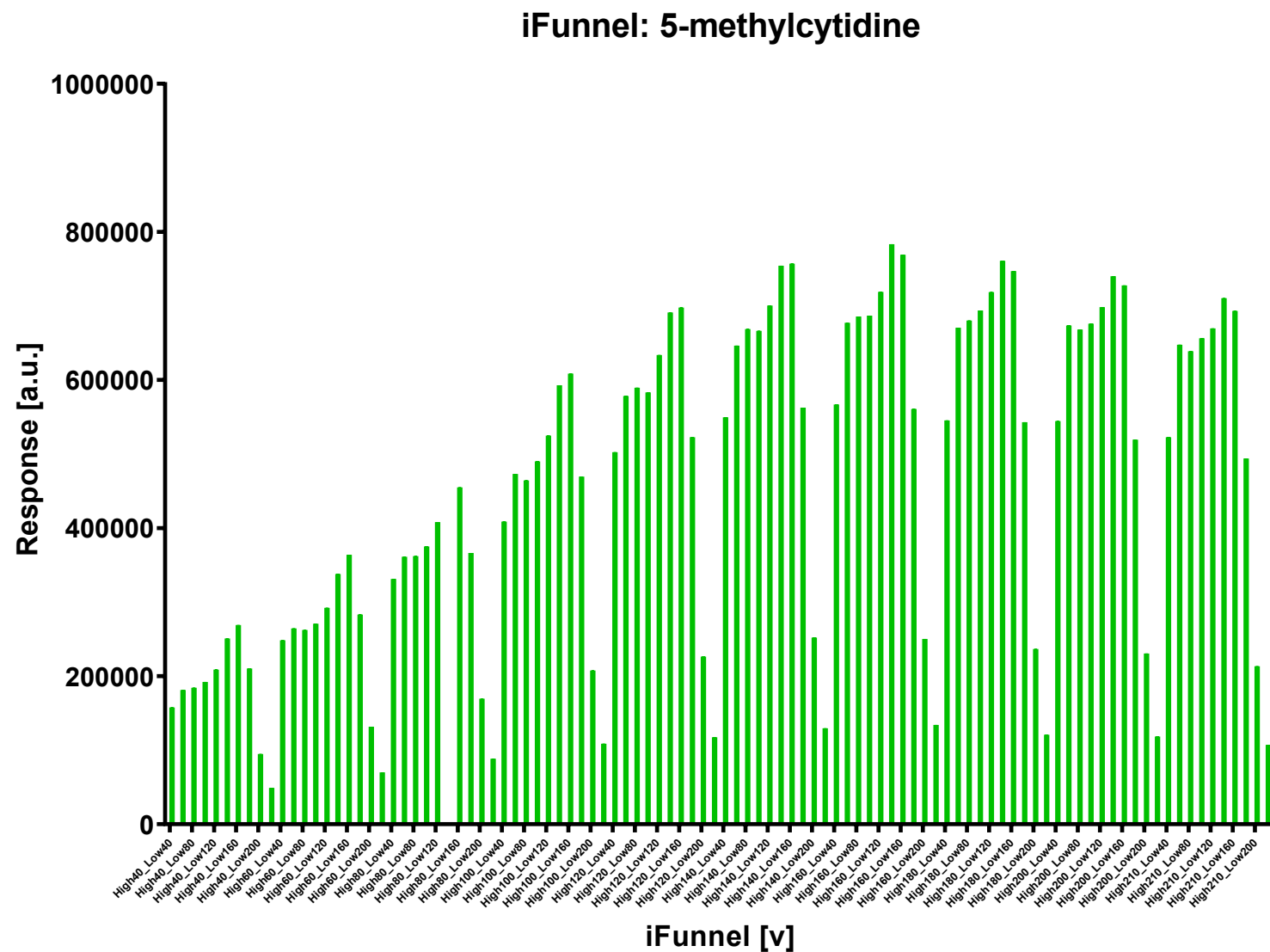


**Fig 3.16** Optimisation of mass spectrometry iFunnel voltages for 5-methylcytidine.

The response values for the iFunnel parameters are shown. The data presented is representative of n=2 technical replicates from the resuspended nucleotides.





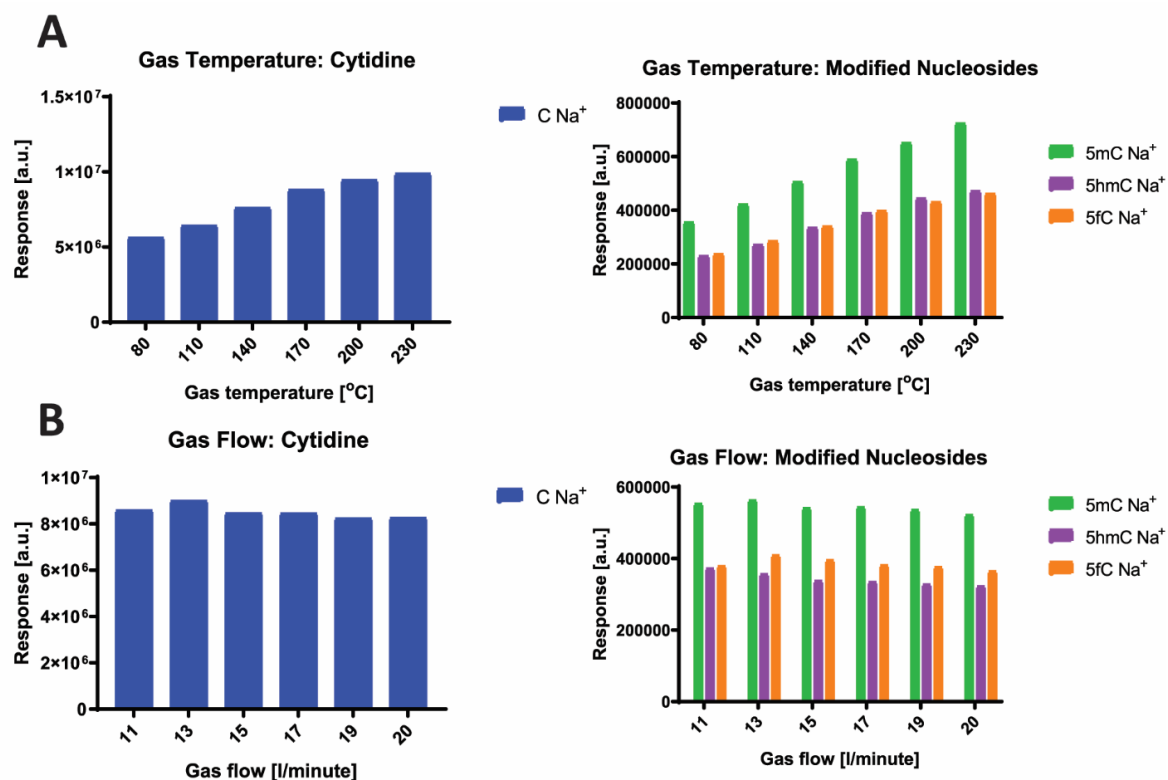


**Fig 3.19 Optimisation of mass spectrometry collision gas.**

The response values for collision gas parameters are shown for cytidine and its modifications. The data presented is representative of n=2 technical replicates from the resuspended nucleotides.

| Parameter                   | Before | Optimised<br>5mC Na <sup>+</sup> | Optimised<br>5hmC Na <sup>+</sup> | Optimal |
|-----------------------------|--------|----------------------------------|-----------------------------------|---------|
| Sheath gas temperature [°C] | 350    | 400                              | 400                               | 350     |
| Sheath gas flow [l/min]     | 12     | 12                               | 12                                | 10      |
| Nozzle [V]                  | 0      | 0                                | 0                                 | 0       |
| Nebulizer pressure [psi]    | 55     | 20                               | 20                                | 20      |
| Capillary [V]               | 3000   | 2500                             | 2500                              | 1500    |
| iFunnel high pressure [V]   | 110    | 160                              | 160                               | 160     |
| iFunnel low pressure [V]    | 60     | 140                              | 160                               | 80      |
| Gas temperature [°C]        | 130    | 230                              | 230                               | 230     |
| Gas flow [l/min]            | 20     | 13                               | 13                                | 11      |

Table 3.3 Optimised mass spectrometry parameters for 5hmC and 5mC detection.

|                      | Original   |            | Optimised  |            |
|----------------------|------------|------------|------------|------------|
| Nucleoside           | LOD [amol] | LOQ [amol] | LOD [amol] | LOQ [amol] |
| 5mC Na <sup>+</sup>  | 25         | 25         | 5          | 25         |
| 5hmC Na <sup>+</sup> | 25         | 50         | 5          | 25         |
| 5fC Na <sup>+</sup>  | 50         | 50         | 50         | 50         |
| C Na <sup>+</sup>    | 500        | 500        | 500        | 500        |

Table 3.4 Limit of detection and limit of quantification for RNA nucleosides before and after mass spectrometry optimisation.

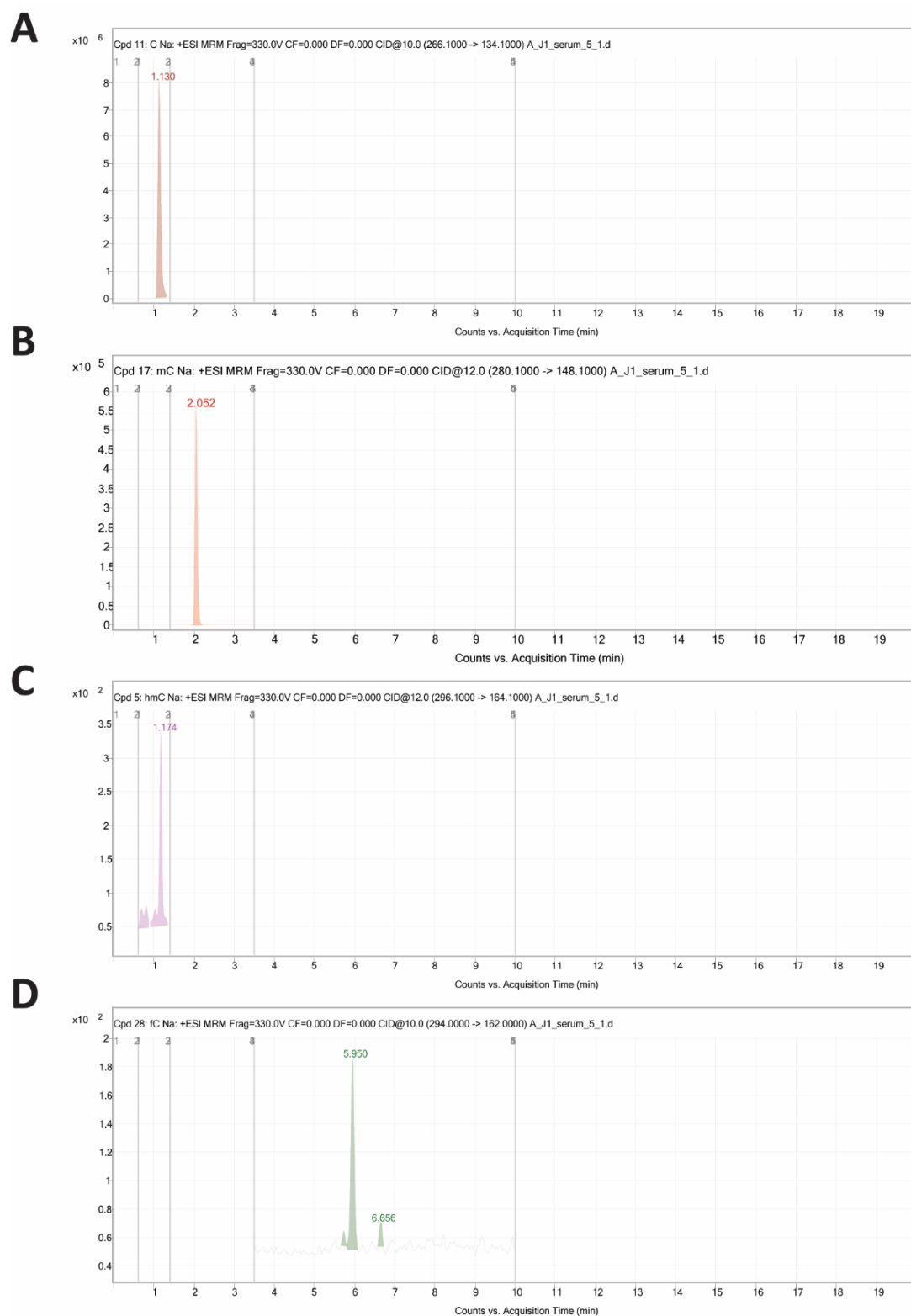


**Fig 3.20 Chromatograms of J1 serum mESC RNA modifications.**

Chromatograms of sodiated cytidine **(A)**, 5-methylcytidine **(B)**, 5-hydroxymethylcytidine **(C)** and 5-formylcytidine **(D)** from 25ng of digested RNA from J1 serum mESCs are shown within their relevant elution time segments. The data presented is representative of n=3 biological replicates.

## 3.4 Discussion

In this chapter I have outlined the work done to improve upon two existing methodologies, each of which displayed their own caveats and limitations beforehand. Most notably, through the use of DNase and RNase A digestions it was shown that RNA dot blotting methodologies accepted within the field are highly susceptible to artefact signals from contaminating DNA. This casts doubt over the validity of published RNA dot blotting data, particularly RNA  $\alpha$ -5mC dot blotting experiments using the common 33D3 clone antibody. When sufficient controls are in place this methodology can be used to generate semi-quantitative data on the abundance of particular RNA modifications.

LC-MS/MS did not display the specificity issues encountered with dot blotting, and from a technical standpoint represented the most promising line of work to develop a highly sensitive method. This was accomplished to a hitherto unachieved level, pushing the technical boundaries of the most sensitive instrumentation available to  $5 \times 10^{-17}$  mol. This provided the ideal basis from which accurate RNA modification quantification could be performed from low cell numbers, such as primordial germ cells.

# Chapter 4: RNA modifications in stem cells

## 4.1 Introduction

In order to evaluate the role(s) of RNA methylation in epigenetic reprogramming, and the transitions that occur within pluripotency leading into differentiation, an *in vitro* stem cell model was used extensively. The infinite proliferative capacity, which defines stem cells, makes them highly amenable to *in vitro* cell culture. This has been a highly fruitful tool in advancing the stem cell field. Embryonic stem cells (ESCs) are derived from the inner cell mass or pre-implantation epiblast of mouse embryos between E3.5-E5.5, whereas epiblast-derived stem cells are generated from the post-implantation epiblast between E5.5-8.0 (Smith, 2017, Osorno et al., 2012, Russ et al., 2000). Media composition and culturing techniques are key to the successful derivation and propagation of any stem cell line. Stem cells can exist in naïve, formative, or primed states, which exist along a developmental axis (Smith, 2017, Kalkan et al., 2017). Naïve cells display low DNA methylation (~2 %), naïve markers (such as *Klf2/4*, *Tbx3*, *Tfcp2l1*), and resist inductive signalling (Leitch et al., 2013, Boroviak et al., 2015, Hayashi et al., 2011, Magnúsdóttir et al., 2013, Kalkan et al., 2017). These cells transition through a formative pluripotent stage, in which early post-implantation markers reflective of the E5.5 epiblast (such as *Otx2*, *Sox3*, *Oct6*) are expressed (Boroviak et al., 2015, Kalkan et al., 2017). In this state the cells are responsive to inductive signalling through FGF and nodal signalling, before becoming primed to differentiate ((Hayashi et al., 2011, Magnúsdóttir et al., 2013, Robertson, 2014, Smith, 2017, Acampora et al., 2016). Primed cells have not yet undergone differentiation, however they are lineage committed and display somatic cell levels of DNA methylation (~4%). ESCs derived in 2i, a media containing 2 inhibitors for MEK (MAP (Mitogen-Activated Protein) Kinase/ERK (Extracellular Signal-Regulated Kinase) Kinase 1) and GSK3 $\beta$  (glycogen synthase kinase 3 beta) respectively, are naïve, whereas EpiSCs derived in epiblast media are considered primed (Ying et al., 2008; Ficz et al., 2013; Habibi et al., 2013, Masaki et al., 2016). ESC derivation in serum-containing media results

in a heterogeneous cell population consisting of a mix of naïve, formative, and primed cells (Smith et al., 2017, Hayashi et al., 2008). Despite this, serum-cultured cell populations display somatic DNA methylation levels (Leitch et al., 2013). Recently it has been shown that formative cells can be successfully isolated from embryos and maintained in *in vitro* culture (Kalkan et al., 2017). Through utilising these stem cell lines, these developmental snapshots can be experimented upon in great detail. This enables the investigation of cell fates spanning the range of pluripotency, leading into primed cells ready to differentiate, as well as a two-fold range of DNA 5mdC abundances.

Additional details of the components and formulations for these media can be found in Chapter 2: Materials and methods section 2.2: Cell culture. All wild type cell lines within the different media conditions outlined above displayed equivalent proliferation rates as observed through careful tissue culture passaging. However, no quantitative proliferation assay was performed on them. Where proliferation rates of the mutant cell lines used deviated from this, is clearly indicated within this body of work.

In addition to establishing and maintaining fixed cell state populations, stem cells can be induced to move along the aforementioned developmental continuum simply by culturing them in the media of the desired cell state (Leitch et al., 2013). Typically, a minimum of 5 passages in the new media is considered sufficient to adopt the new cell state, although further passaging in 2i is necessary to erase DNA methylation at imprint control regions (ICRs) (Leitch et al., 2013, Choi et al., 2017, Yagi et al., 2017). However, morphological changes are identifiable over a much more rapid timescale. By changing culturing conditions, DNA methylation levels can thus be altered from somatic levels of ~4% of all cytidines to ~2%. This has been shown to be driven by the inhibition of the de novo DNA methyltransferases *Dnmt3a* and *Dnmt3b* by *Prdm14*, as has been found *in vivo* (Leitch et al., 2013, Yamaji et al., 2008). More drastic perturbation can be achieved upon *Dnmt1*, *Dnmt3a*, and *Dnmt3b* triple knockout (TKO), which results in no

detectable DNA methylation regardless of culturing conditions (Tsumura et al., 2006, Ficz et al., 2011, Williams et al., 2011). Through combining these systems, DNA methylation can be modulated from somatic cell maximum levels down to complete ablation. As such, RNA modifications and RNA methyltransferases (RNMTs) can be assayed within an *in vitro* DNA 5mdC demethylation system, along with states across a naïve-primed pluripotency axis.

The metabolisms of DNA 5mdC and RNA 5mC both involve the action of the ten eleven translocation (TET) enzymes, which successively oxidise the methyl group to hydroxymethyl-, formyl-, and carboxy(deoxy)cytosine (Tahiliani et al., 2009; Ito et al., 2011; He et al., 2011, Fu et al., 2014). TET1<sup>-/-</sup>, TET2<sup>-/-</sup>, and TET3<sup>-/-</sup> triple knockout (TTKO) ESCs were used to provide a cell line devoid to any TET enzyme activity with which to assess how this may affect RNA 5mC and 5hmC levels.

For further background, please see Introduction and Literature Review section 1.2

## 4.2 Experimental Aims

Within this chapter stem cells were used to model DNA demethylation and epigenetic reprogramming, as well as developmental progression from naïve to primed fates. In order to achieve the full repertoire of cell states desired, *Dnmt1*, *Dnmt3a*, and *Dnmt3b* triple knockout (TKO) epiblast-derived-like cells (TKO EpiLSCs) were generated and validated. It was unclear before beginning this work if it would be possible to generate an epiblast-like state without the canonical DNA methyltransferases. RNMTs were analysed in regards to their expression and subcellular localisation in an attempt to understand their regulation and functionality during these cell state transitions. Additionally, RNA modification levels were quantified in both total RNA and within RNA subpopulations.

## 4.3 Results

### 4.3.1 Descriptions and derivations of cell lines

A number of different pluripotent stem cell states are analysed within this body of work. While the rationale around the need for these stem cell fates is given in the above introduction, it is worth noting the origins and backgrounds of the specific cell lines used.

#### 4.3.1.1 Wild types

The original source of each cell line is given in Materials and Methods Table 2.2. Wild type (WT) ESCs of the 129 background were used in the form of J1 cells. Additionally, E14 cells, of the same genetic background were used. Within the 129 background these cells were originally derived from embryos of differing sub-strain, with J1 cells being 129/SvJ while E14 ESCs are 129/Ola. The sub-strain differences are negligible and pale in comparison to the variation within natural mouse populations in the wild. These two cell lines were used as wild type controls in different experimental branches. Both have been extensively worked with in the stem cell field, however E14 cells are not truly wild type. It is often overlooked that they are Hypoxanthine phosphoribosyl-transferase (HPRT) knockouts, generated through retrotransposon insertion into the third exon (Tsuda et al., 1997). HPRT is a gene involved in the nucleotide salvage pathway, which historically was knocked out to better facilitate marker selection of any additional genetic manipulations of these stem cells (Tsuda et al., 1997). This increased the success rate in generating knockout mouse lines, which has played a major role in the ubiquity of the E14 ESC cell line. Intellectually, both J1 and E14 cells sufficiently represented wild type cell states for my experiments and were cultured in both serum and 2i media.

EpiSC were also derived from the 129 background and were graciously received from Prof. Austin Smith's lab via Dr. Carla Mulas and Dr. Harry Leitch.

#### 4.3.1.2 DNMT and TET triple knockouts ESCs

DNMT1<sup>-/-</sup>, DNMT3A<sup>-/-</sup>, and DNMT3B<sup>-/-</sup> combinatorial knockouts were generated from J1 cells (Tsumura et al., 2006). Initially DNMT3A<sup>-/-</sup>, and DNMT3B<sup>-/-</sup> double knockout cells were generated by deleting motifs IV and VI from the catalytic domain of each gene (Okano et al., 1999). To this cell line exons 31-33 of DNMT1 were removed to give the final DNMT triple knockout (TKO) cell line (Tsumura et al., 2006). J1 cells also served as the starting point for the formation of TET1<sup>-/-</sup>, TET2<sup>-/-</sup>, and TET3<sup>-/-</sup> triple knockout (TTKO) cells by CRISPR (Hill et al., 2018). These cells were generated and kindly provided Dr. Cristina Requena Torres. Both cell lines were cultured independently in serum and 2i media.

#### 4.3.1.3 EpiLSC TKO generation and validation

DNA methylation is necessary for differentiation to occur. Attempts to force DNMT TKO ESCs to differentiate through altering media compensation result in apoptosis (Schmidt et al., 2012). TKO cells can be stably propagated as pluripotent stem cells; however, it was unclear if they were capable of adopting a primed cell fate. In order to investigate this and potentially enable the comparison of EpiSCs lacking DNA 5mdC with wild type EpiSCs, TKO cells were cultured in EpiSC media. Figure 4.1 shows the morphology of EpiSCs, as well as ESCs cultured in serum and 2i. The TKO cells appear as their wild type equivalents grown in the same media. The TKO cells in EpiSC media were termed DNMT triple knockout epiblast-derived-like stem cells (TKO EpiLSCs), as these cells were not originally derived from epiblast tissue. As such, it would have been a misnomer to merely name them TKO EpiSCs.

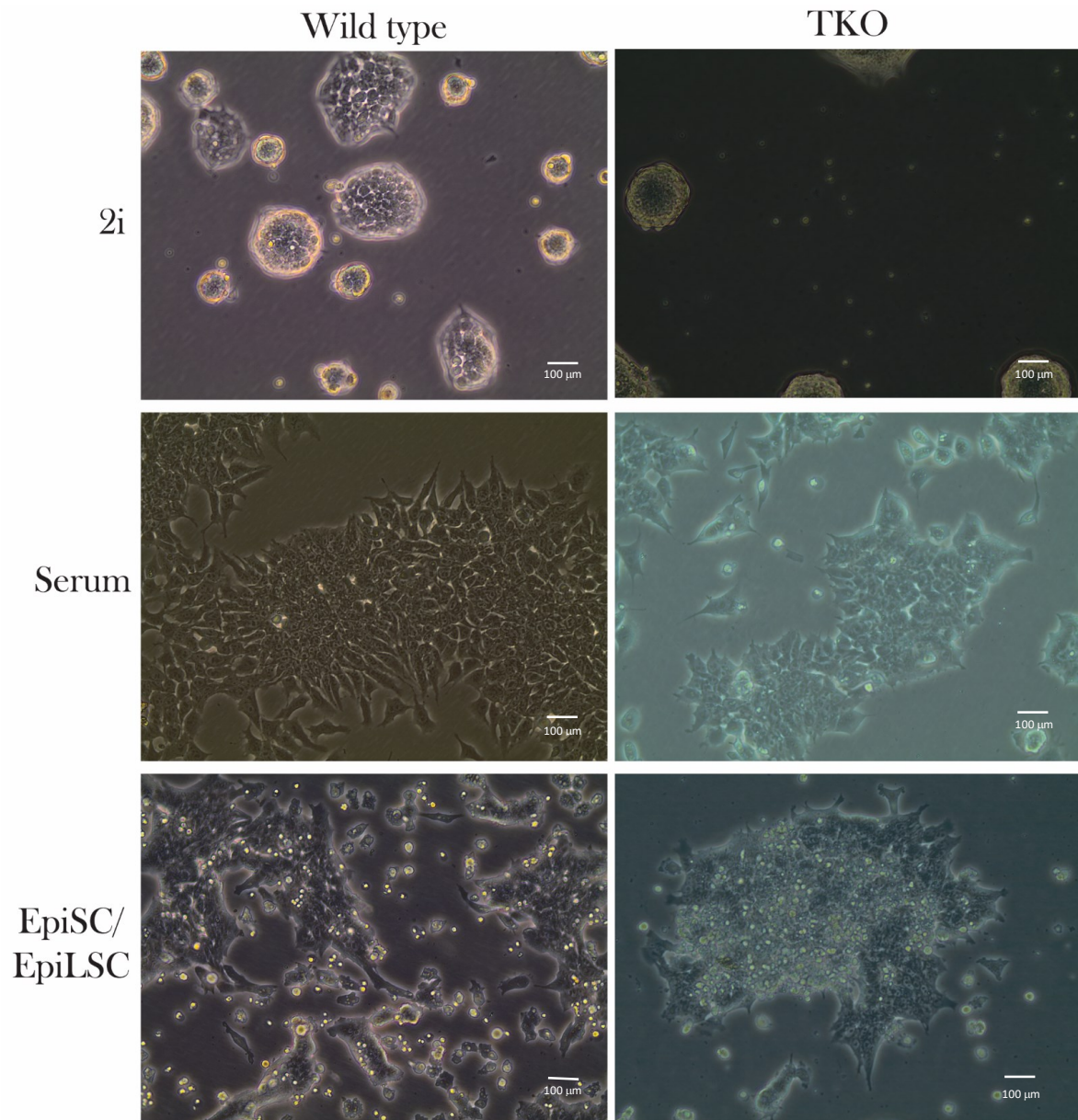


**Fig 4.1 Brightfield images of TKO and wild type stem cells in 2i, serum, or EpiSC conditions.**

Brightfield images of the cell lines used in the *in vitro* reprogramming model are displayed. E14s are displayed as wild type ESCs. Wild type EpiSCs are shown, while the TKO cells in an Epiblast-derived stem cell state are epi-like stem cells (TKO EpiLSCs). This is a result of their generation from TKO embryonic stem cells cultured in EpiSC media. All cell lines displayed are self-renewing. Images are reproduced at a magnification of 44x with the scale bar displaying 100  $\mu\text{m}$ . Cell lines were observed daily while in culture. The images displayed are representative of these cell lines.

In order to more fully characterise this novel cell line, RT-qPCR was performed for differing lineage markers upon RNA extracted from TKO EpiLSC, genuine EpiSC, and TKO serum cells. TKO serum cells acted as a negative control, as it was from these cells that the TKO EpiLSCs originated (Fig. 4.2 and Fig. 4.3). The pre-implantation epiblast markers reduced expression 1 (*Rex1*) and *Nr0b1* are highly downregulated in EpiSCs and TKO EpiLSCs, while the EpiSC and primitive ectoderm marker fibroblast growth factor 5 (*Fgf5*) is substantially upregulated (Fig. 4.2A & B) (Kim et al., 2013). This is consistent with TKO EpiLSC transitioning away from an ES cell state and successfully adopting an EpiSC fate. However, as *Rex1* was not downregulated to the same extent as in EpiSCs, the entire cell population may not have fully transitioned. To determine if any further lineages had become specified, additional lineage markers were assayed. No upregulation of trophectoderm or primitive endoderm lineages was observed for TKO EpiLSCs. This was despite EpiSCs displaying an unexpected increase in Sry-related HMG box gene-17 (*Sox17*) expression relative to serum-cultured ESCs (Fig. 4.3A & B) (Chrysanthou et al., 2018, Artus et al., 2011). Epithelial and mesoderm markers displayed divergence between the markers used. This manifested in the upregulation of Cadherin-2 (*Chd2*) but not vimentin (*Vim*) and in *Mus musculus Brachyury (T)* but not even-skipped homeobox 1 (*Evx1*) or mesoderm posterior 1 (*Mesp1*) (Fig. 4.3C & E) (Chrysanthou et al., 2018, Wen et al., 2017). It was thus unclear if TKO EpiLSCs were progressing down these lineages, although it was reassuring to see the same expression levels in EpiSCs. Accordingly, this was interpreted as being a reflection of the EpiSC state. Figure 4.3C shows that all three epithelial-mesenchymal transition (EMT) markers used are more highly expressed in TKO EpiLSCs than EpiSCs (Chrysanthou et al., 2018). The magnitudes of the expression differences are minor, and for Snail Family Transcriptional Repressor 1 and 2 (*Snai1* and *Snai2*) are mostly due to EpiSC exhibiting lower levels than ES cells grown in serum. As such, it is only zinc finger E-box binding homeobox 2 (*Zeb2*) for which a true upregulation of expression is found in TKO cells upon

culturing in EpiSC conditions. EMT is a process by which epithelial cells reduce adhesion to one another to enable migration and invasion of other cell layers. *In vivo*, an EMT from post-implantation epiblast cells would correlate to gastrulation. Taken together, it appeared that the TKO EpiLSCs mostly existed within the EpiSC state, although some pre-implantation epiblast character was retained while some progression into an epithelial-mesenchymal transition also occurred. It may have been intriguing to address whether the modest pre-implantation epiblast and EMT signals were a result of distinct minor cell populations within a main population of more clearly definable EpiSC-like cells. However, from the results given it was considered that these cells display sufficient similarity to EpiSCs to continue with for RNMT and RNA modification analysis.

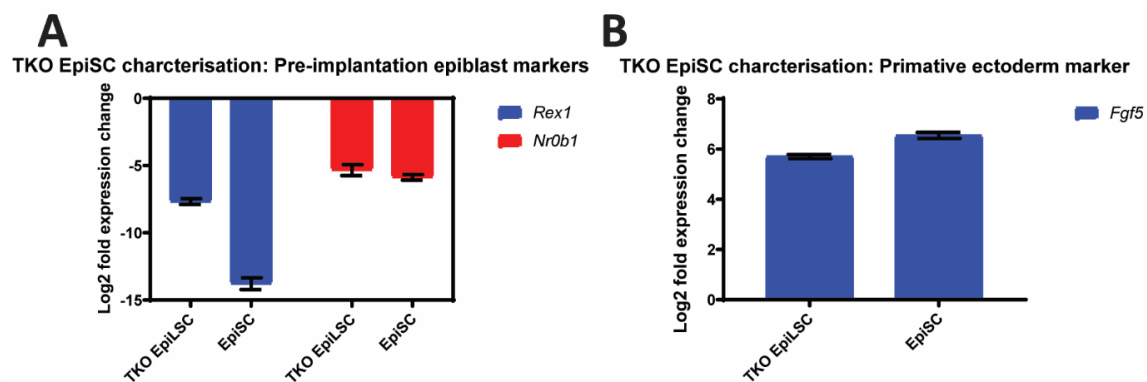


**Fig 4.2** TKO EpiLSCs transition from a ESC fate to adopt a EpiSCs fate as determined by RT-qPCR.

RT-qPCR was performed on RNA extracted from TKO EpiLSCs EpiSC and TKO serum cells. The log fold change relative to TKO serum cells is provided for the TKO EpiLSCs under investigation, as well as the positive control EpiSCs, for markers of **(A)** the pre-implantation epiblast and **(B)** primitive ectoderm. Error bars give standard deviation (n=3).

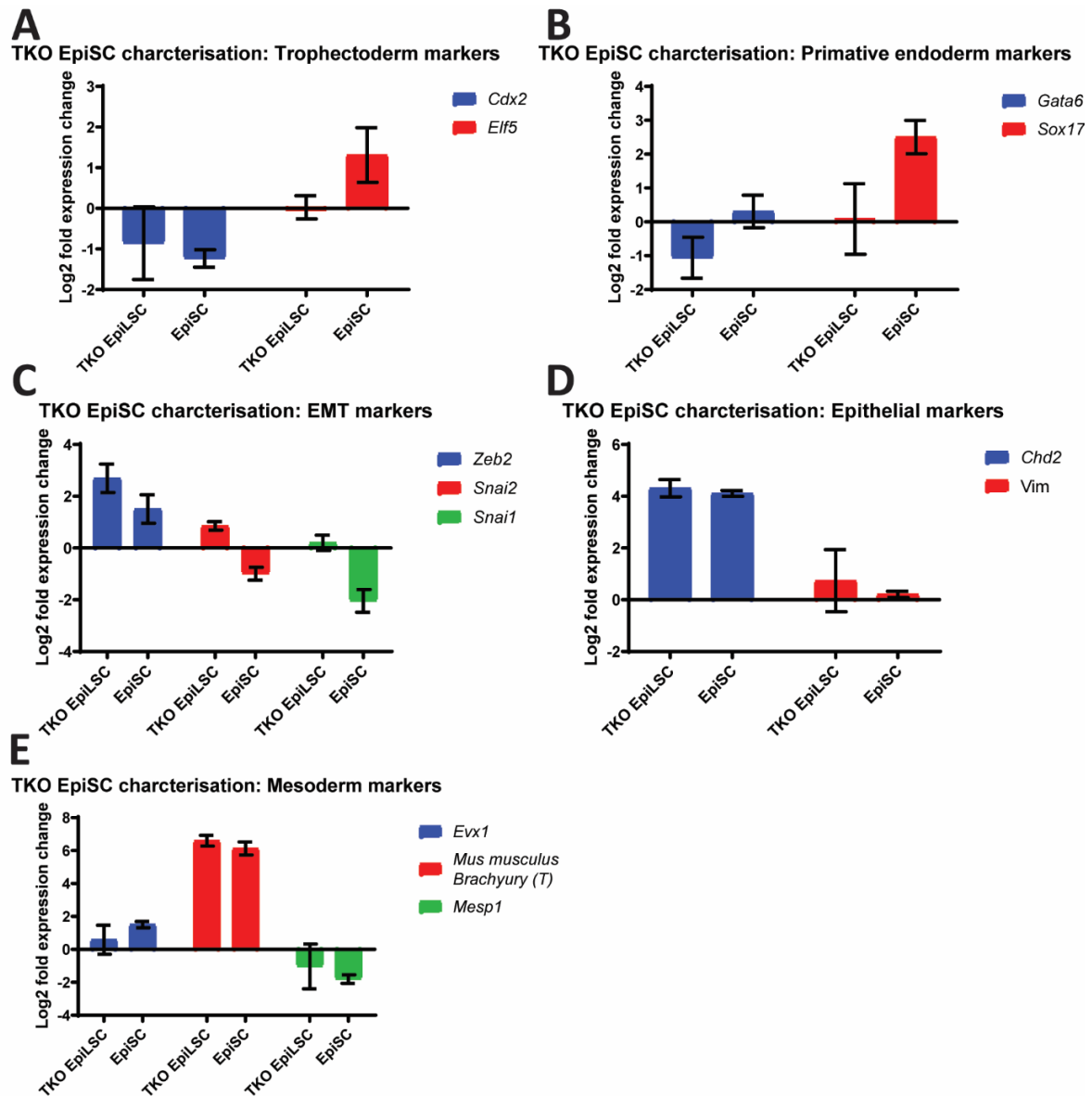


**Fig 4.3** TKO EpiLSCs initiate some programs indicating transition beyond the EpiSC fate.

**(A-E)** RT-qPCR was performed on RNA extracted from TKO EpiLSCs EpiSC and TKO serum cells. The log fold change relative to TKO serum cells is provided for the TKO EpiLSCs under investigation, as well as the positive control EpiSCs, for markers of assorted lineages. **(A)** Trophectoderm, **(B)** primitive endoderm, **(C)** EMT, **(D)** epithelial, and **(E)** mesoderm markers were assessed. Error bars give standard deviation (n=3).

### 4.3.2 RNMT gene expression

#### 4.3.2.1 mRNA quantification reveals increased expression of multiple RNMTs in DNA hypomethylated cells

RT-qPCR was performed on RNA from each cell state depicted in Figure 4.1 and analysed alongside previously published microarray expression data (Leitch et al., 2013). The expressions of the 8 RNMTs encoded in the mammalian genome were assayed using the microarray data. Having performed expression analysis in *in vivo* PGC reprogramming (*In Vivo* Work Figure 5.1) NOP2/Sun domain protein 7 (NSUN7) was detected at such a low level *in vivo*, as well as *in vitro* by microarray, that it was not considered for RT-qPCR. GAPDH was used as a housekeeping gene for RT-qPCR.

The effects of culturing media were considered first. Culturing cells in 2i media compared to serum, and the concurrent transition to a fully naïve cell state with a hypomethylated genome, results in minimal changes in RNMT expression at the mRNA level. Figure 4.4A and B differ only in the methodology used to measure mRNA expression, but in spite of this give differing expression change profiles. Both techniques give no change in expression of *Nop2* and a modest downregulation in *Nsun2*. *Trdmt1*, *Nsun3*, and *Nsun6* are all upregulated by RT-qPCR, but only *Nsun3* shows any upregulation by microarray. *Nsun4* and *Nsun5* show some upregulation and downregulation respectively by microarray, but these results are not consistent with the RT-qPCR results. The magnitude of any given expression change is low (typically around  $\log_2 0.5$ ), but the combinatorial effect of these RNA methyltransferase gene expression changes could give a more pronounced alteration on RNA 5mC levels. Culturing TKO ESCs in 2i media results in no changes in RNMT expression at the RNA level, indicating that the minor changes observed in wild type cells are ablated by the knockout of the canonical DNMTs (Figure 4.4C).

The transition to a primed cell state in EpiSCs is accompanied by the reduction of *Nsun2* and *Nsun4* expression, which is observed in both wild type and TKO cells (Figure 4.4D and E). Conversely, *Trdm1* and *Nsun5* are upregulated, although only in wild type cells. As with the transition to 2i, deviations are observed in the expression profiles depending upon whether wild type or TKO cells are assessed. TKO EpiLSCs display a more repressed profile for all RNMTs except *Nsun6* when culturing conditions are isolated compared to wild type cells.

Analysing the effect DNMT triple knockout has on gene expression independent of culturing conditions allows for the disentangling of phenomena brought about by DNA methylation from those which arise from cells shifting between naïve, formative, and primed states. In cells grown in 2i media *Nop2* and *Nsun5* stand out as being upregulated, upon DNMT TKO, while in serum media *Trdm1*, *Nop2*, *Nsun2*, *Nsun3*, and *Nsun5* are all upregulated (Figure 4.5A and B). Curiously, no such changes are displayed in TKO EpiLSCs when compared to EpiSCs (Figure 4.5C). That the RNMT expression changes brought about by removing DNA 5mdC are not recapitulated in primed cells highlights naïve and formative pluripotency as being pivotal to understanding RNMT functionality.

By analysing the RNMT mRNA expression within this *in vitro* reprogramming and priming model some exciting results emerged. However, as it is at the protein level that these gene products are active the protein expression was assayed by western blotting.

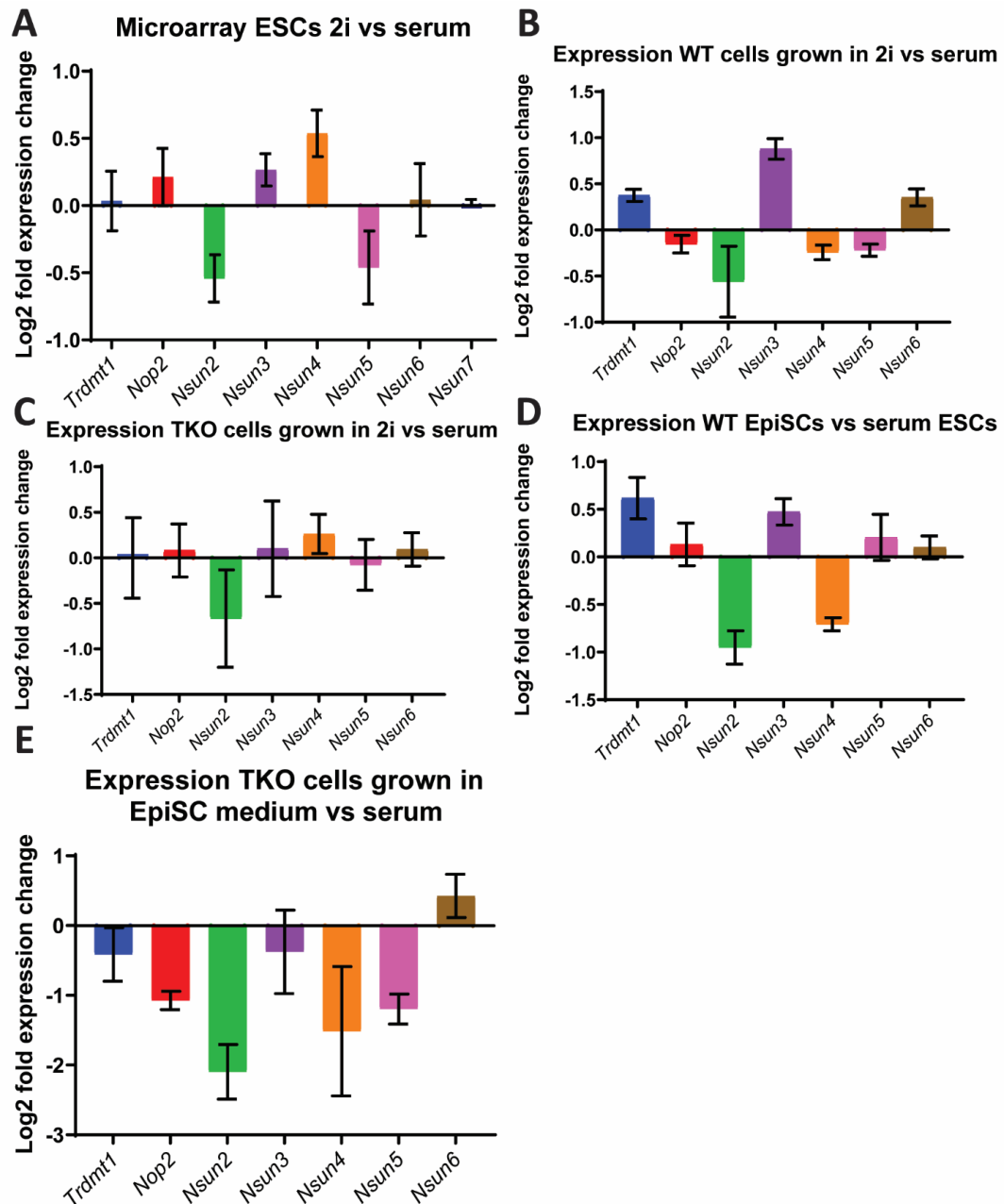


**Fig 4.4 RNMT mRNA expression level in different pluripotent stem cell fates as measured by microarray and RT-qPCR reveals subtitle variations in expression.**

**(A)** Previously published microarray data was used to quantify the expression change of mammalian RNMTS in mESCs cultured in 2i media conditions compared to serum (Leitch et al., 2013). Error bars give standard deviation (n=5). **(B-E)** RT-qPCR was performed on RNA extracted from ESCs grown in 2i or serum or EpiSC, along with their TKO equivalents. E14s were used as wild type ESCs. GAPDH was used as a housekeeping gene for RT-qPCR. The log fold changes for 2i **(B-C)** and EpiSC **(D-E)** are given. Error bars give standard deviation (n=3).

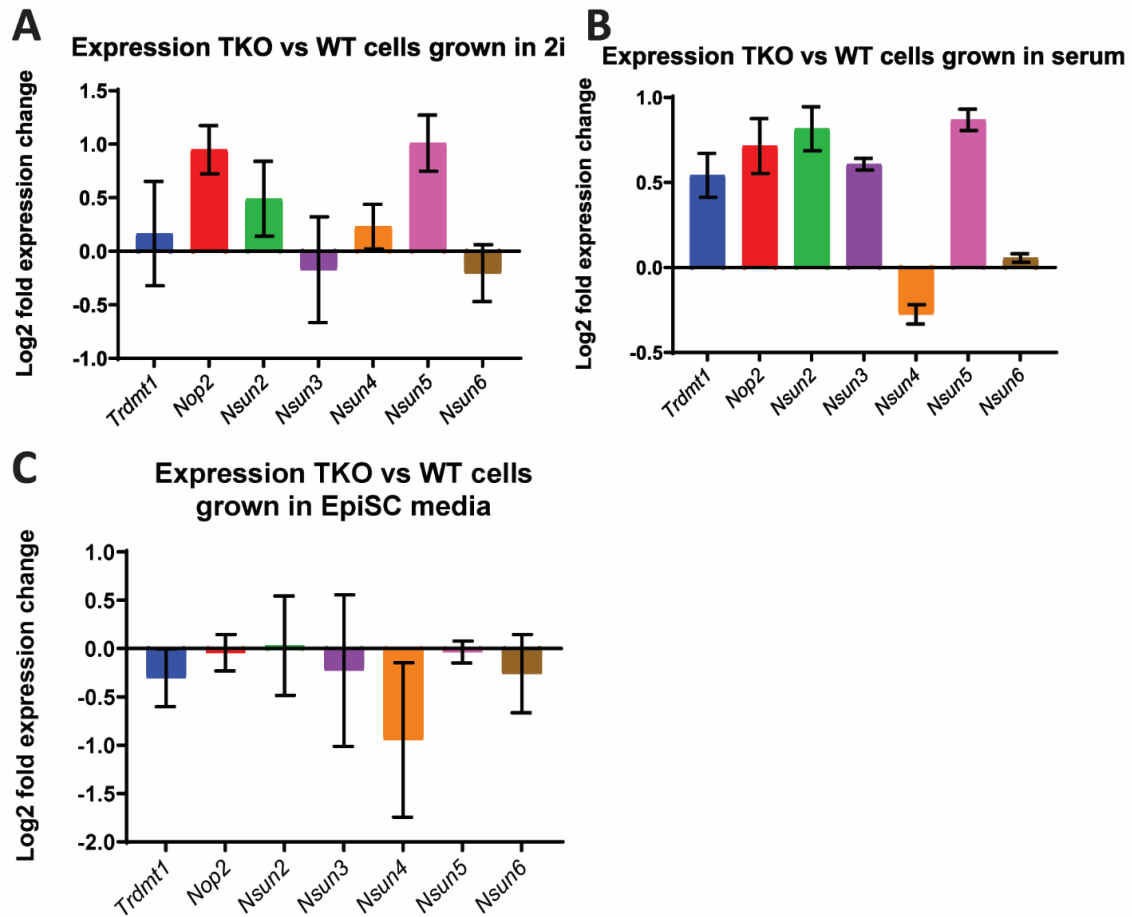


**Fig 4.5 RNMT mRNA expression level upon ablation of DNA methylation as measured by RT-qPCR reveals increased expression of multiple RNMTs.**

RT-qPCR was performed on RNA extracted from ESCs grown in 2i or serum or EpiSC, along with their TKO equivalents. E14s were used as wild type ESCs. GAPDH was used as a housekeeping gene for RT-qPCR. The log fold changes for DNMT triple-knockout in cells cultured in 2i (**A**), serum (**B**), or EpiSC (**C**) media are given. Error bars give standard deviation (n=3).

#### 4.3.2.2 Protein abundance reveals increased NOP2, NSUN2 and TRDMT1 expression in DNA hypomethylated, non-naïve cells

From the mRNA expression data depicted in Figures 4.4 and 4.5, as well as expression data from primordial germ cells (In Vivo Work Figure 5.1), NOP2, NSUN2, NSUN5 and TRDMT1 were selected as promising candidates for further investigation. In order to determine the expression level of these proteins, western blotting was performed on wild type and TKO cells cultured in serum and 2i conditions, with representative blots provided in Figure 4.6. This represents the *in vitro* reprogramming system displaying DNA 5mdC levels at somatic cell levels of 4% (WT

serum), hypomethylated (WT 2i) 2% and entirely demethylated (TKO cells). Total histone H3 was selected as a loading control and detected using a pan-H3 antibody whose binding is independent of histone modification status. NOP2, NSUN2, and TRDMT1 are all most highly expressed in the TKO serum cell state (Figure 4.6A, C, and D). While wild type cells and TKO 2i cells display similar levels of NOP2 and TRDMT to one another, 2i culture reduces the expression of NSUN2 in both wild type and mutant cell lines. NSUN2 displays a double-banding pattern which is only observed in TKO cells. NSUN2 is predicted to be highly modified, including phosphorylation, acetylation, malonylation and glycylation. Some of these modifications have been verified experimentally (Sakita-Suto et al., 2007). This double-banding can thus be explained by two distinct populations of differently modified enzymes, which is highly indicative of regulation. Either 2i culture or DNMT TKO cause a reduction in NSUN5 expression, with the combination of the two reducing protein expression down to barely detectable levels in the TKO 2i cell state.

It thus appears that in the absence of canonical DNA methyltransferase activity and DNA 5mdC, NOP2, NSUN2, and TRDMT1 are all upregulated in a heterogeneous cell population, but not in a solely naïve one. NSUN5 is reduced in its expression in the absence of 5mdC in a manner that is magnified, not ablated by the naivety of ESCs cultured in 2i. This shows two disparate mechanisms by which RNMTs respond to the loss of DNA methylation, which, in the case of NSUN2, are linked to posttranslational modification(s). It is worth noting that in primordial germ cells 5mdC levels fall to around 0.3% at the nadir of reprogramming, and so the 2% methylation in wild type cells grown in 2i media may not be sufficient to trigger this response (Hill et al., 2018). That both upregulation and downregulation are displayed by the RNMTs assayed by western blot highlights the capacity of the differing RNMTs to function independently and validated a forward screening approach. Naïve ESCs in 2i media show reduced levels of NSUN2 and NSUN5, showing a DNA methylation-independent pathway to regulate these genes.

The upregulation of NOP2, NSUN2, and TRDMT1 upon DNMT TKO in serum conditions is consistent at both the mRNA and protein level, which was encouraging. There are some discrepancies between the expression patterns found by western blotting and those observed at the mRNA transcript level. This is to be expected, as it has been shown that mRNA abundance only explains 79 % of corresponding protein expression variation in yeast (Lu et al., 2007, Ramakrishnan et al., 2009). The lack of correlation between mRNA and protein abundance in particular instances can be used to postulate regulation mechanisms involving translational efficiency or decay rates. The clearest example of this is NSUN5, which at the mRNA transcript level is upregulated upon DNMT TKO, while the protein is less abundant in this cell state. Lowered translational efficiency and/or increased protein degradation could mechanistically explain this. Having assessed the expression dynamics for candidate RNMTs, their subcellular localisation was assessed to further probe their functionality and potential regulation.

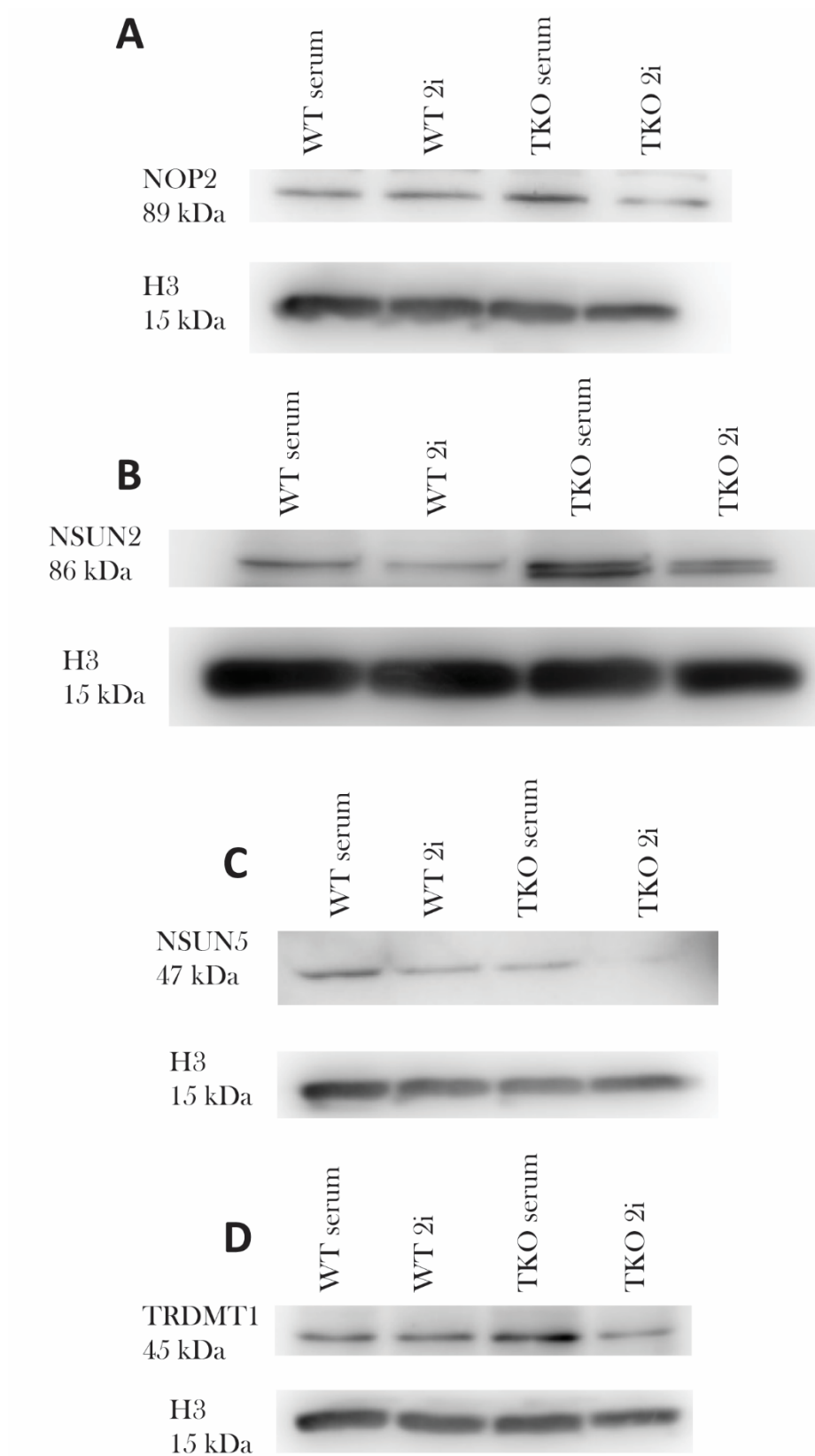


**Fig 4.6 RNMT protein expression level as measured by western blotting reveals dynamic expression changes.**

(A) Western blotting of (A) NOP2, (B) NSUN2, (C) NSUN5 and (D) TRDMT1 from wild type and TKO ESCs cultured in 2i or serum is displayed. Total histone H3 was selected as a loading control. Representative blots are displayed of (A) n=3, (B) n=3, (C) n=3 or (D) n=2 biological replicates. J1s were used as wild type ESCs.

### 4.3.3 RNMT subcellular localisation reveals novel NOP2 and NSUN2 distributions

RNA processing reactions most commonly occur in the nucleus, primarily within the nucleolus (Hernandez-Verdun et al., 2010). As such, subcellular localisation can be used to infer functionality and regulation for RNMTs. NOP2 was initially discovered as a nuclear-associated protein and while NSUN2 has been observed to undergo a cell cycle stage-dependent cycling between the nucleolus and cytoplasm (Freeman et al., 1988, Sakita-Suto et al., 2007). NSUN4 is restricted to the mitochondria and TRDMT1 has been observed in the cytoplasm (Metodiev et al, 2014, Lewinska et al., 2015). The other RNMTs are predicted to be nucleolar, but this had not been experimentally verified.

Immunofluorescent staining for NOP2, NSUN2, NSUN5 and TRDMT was performed upon stem cell lines, with representative images provided in Figures 4.7-4.10. In each experiment DAPI was used to stain DNA, while immunofluorescent staining of octamer-binding transcription factor 4 (OCT4) provided a nuclear marker which is excluded from the nucleolus. Initially staining was performed upon wild type and TKO ESCs grown on coverslips in serum and 2i conditions. The stainings were performed in parallel and coverslips placed upon the same microscope slides. Immunofluorescent staining was also performed using EpiSC and TKO EpiLSCs, however due to spatial requirements coverslips with these cells grown upon them could not be placed on the same slides as the other cells. As the same imaging parameters can be used across any given slide, this enabled the expression of RNMTs to be compared within the ESC sample set, but not with EpiSC or EpiLSCs. This was in addition to assaying the subcellular localisation of the RNMTs, for which all stem cell states could be validly compared.

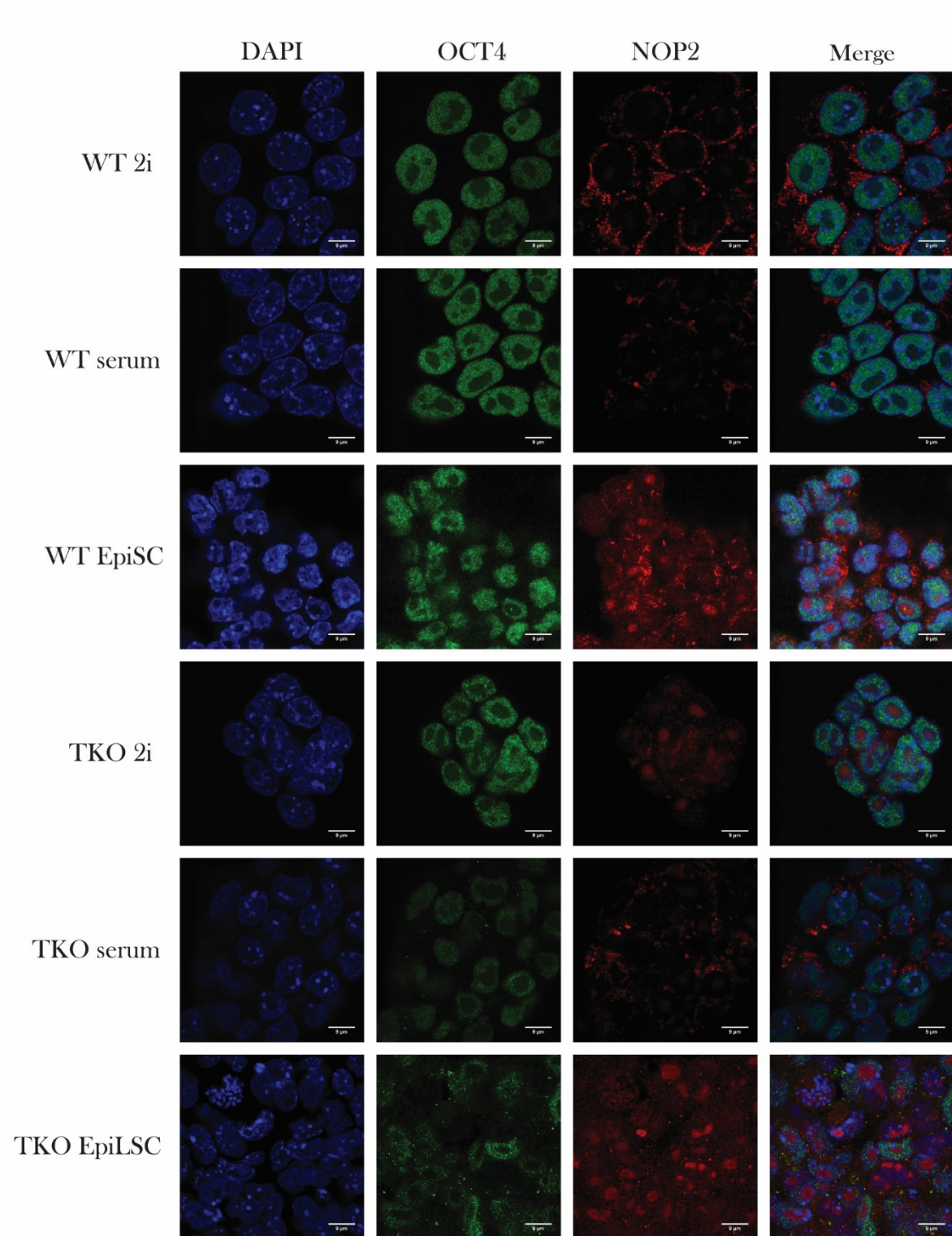


**Fig 4.7 NOP2 subcellular localisation.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu\text{m}$ .

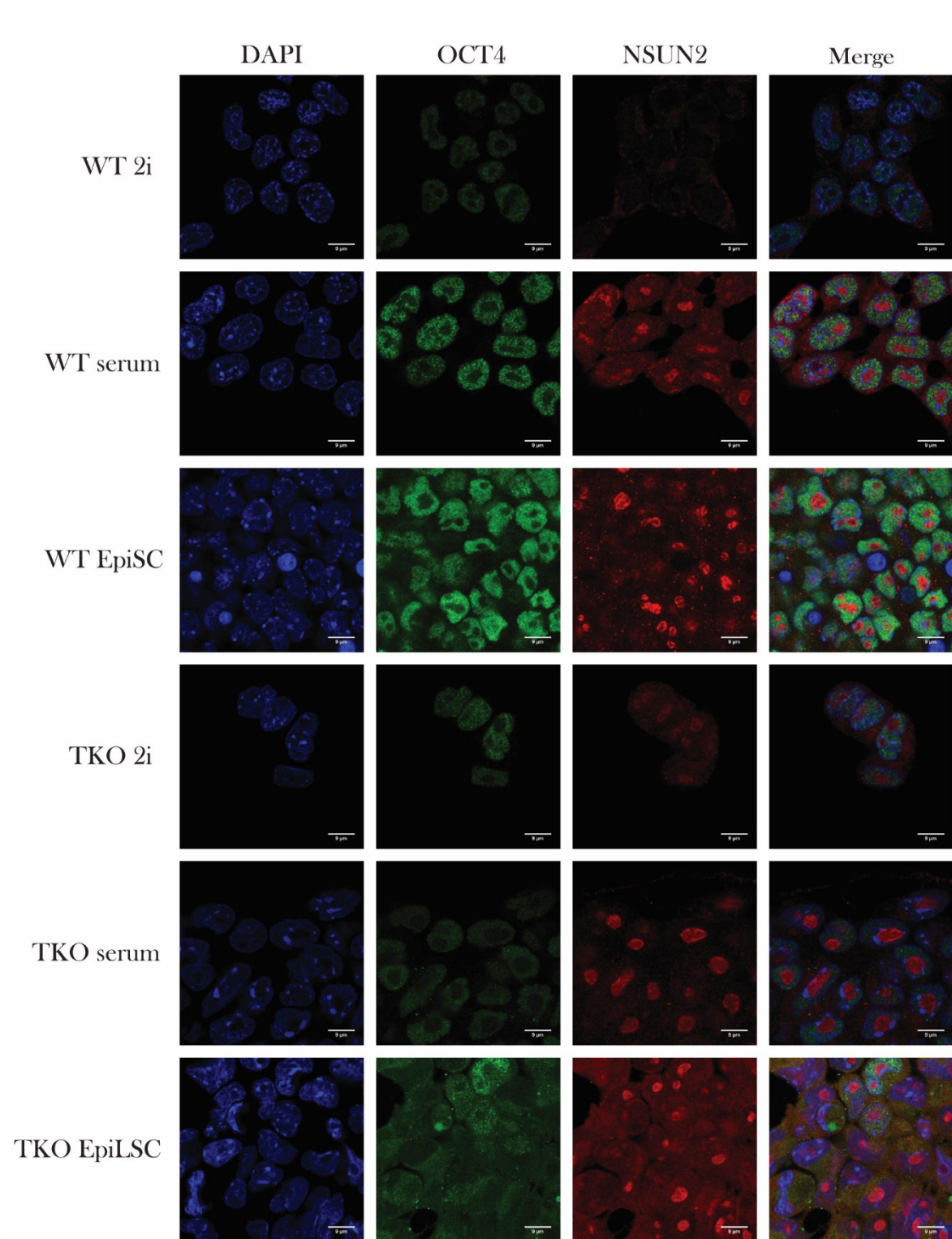


**Fig 4.8 NSUN2 subcellular localisation.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu\text{m}$ .

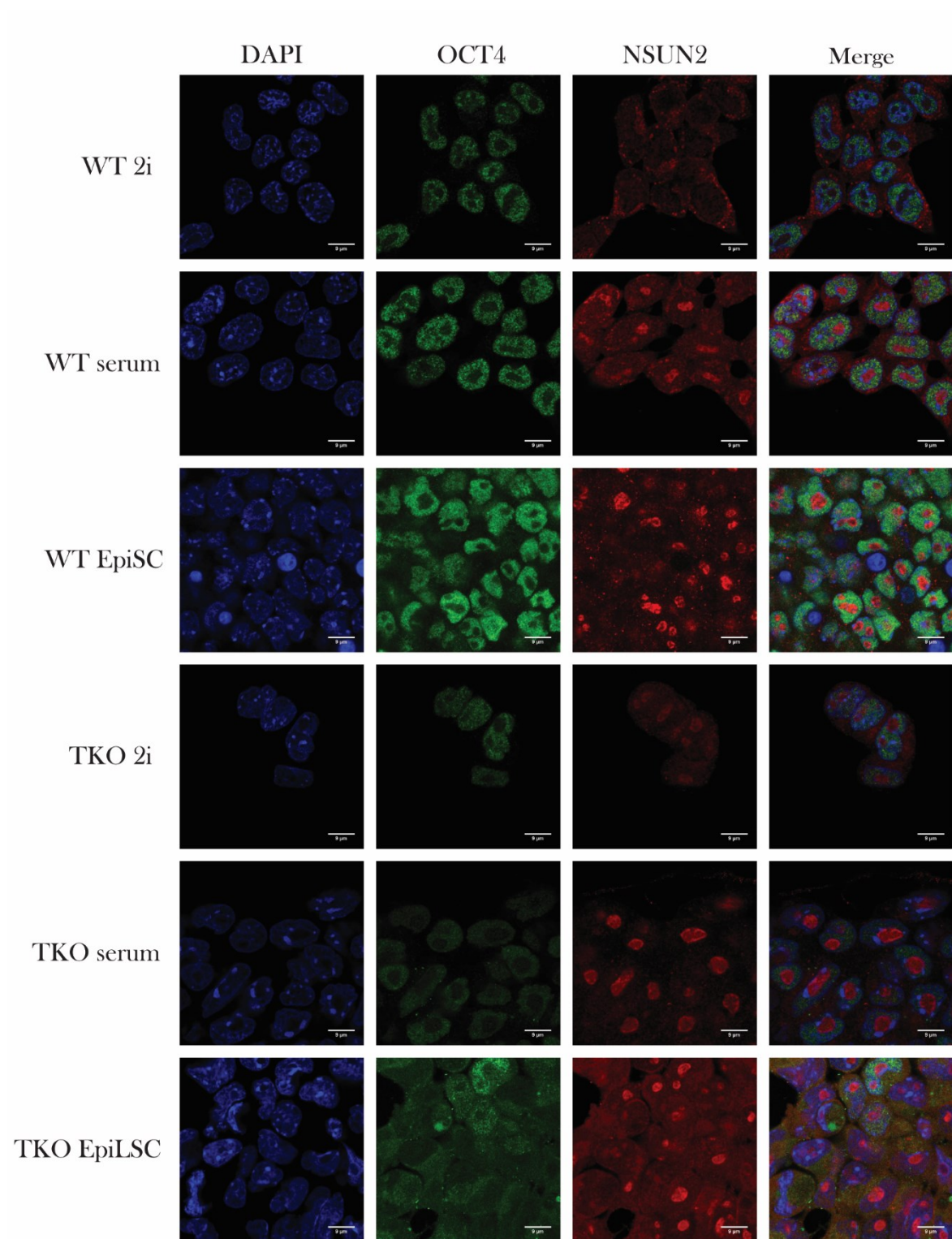


**Fig 4.9 NSUN2 subcellular localisation – WT 2i signal boosted.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu\text{m}$ .

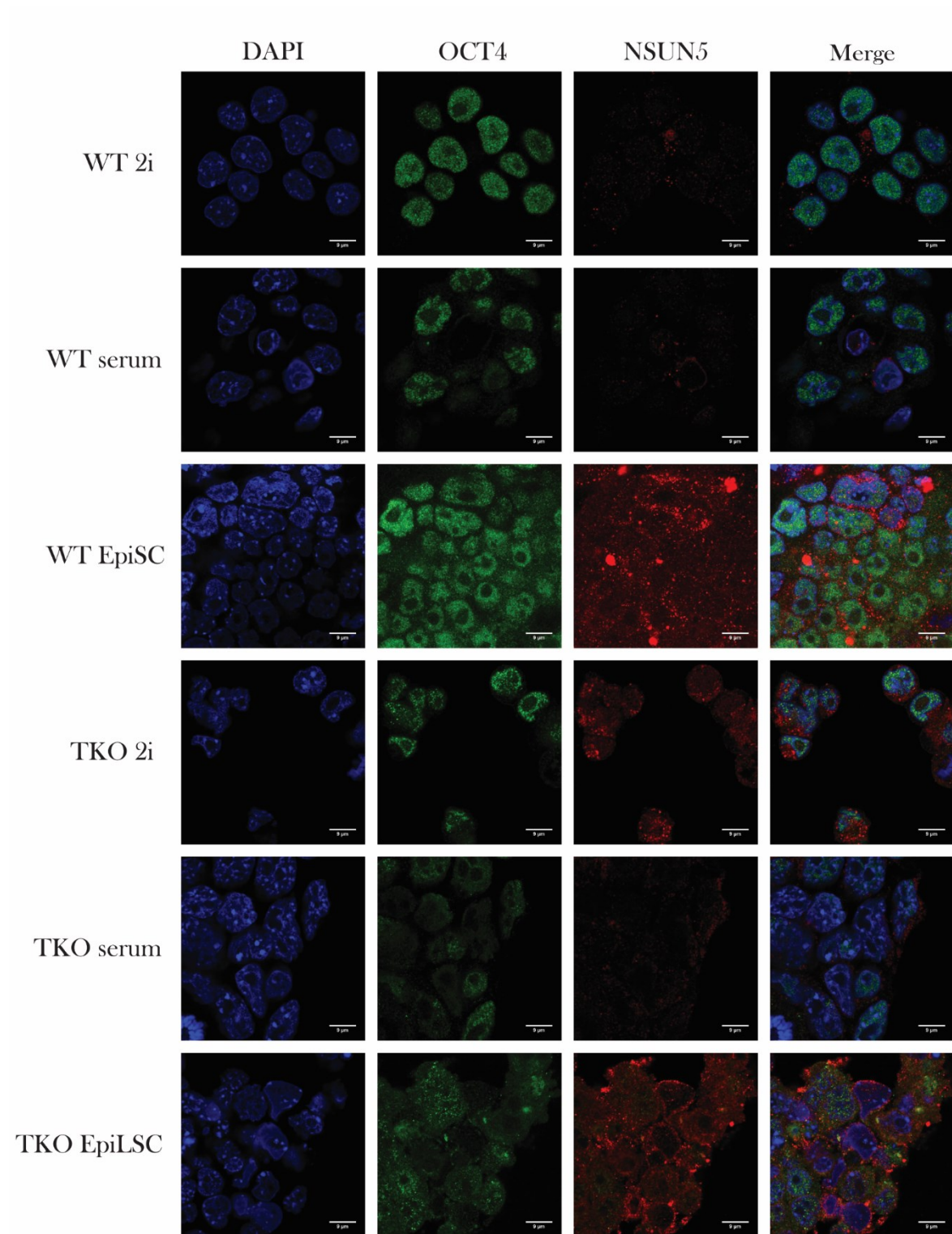


**Fig 4.10 NSUN5 subcellular localisation.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu\text{m}$ .

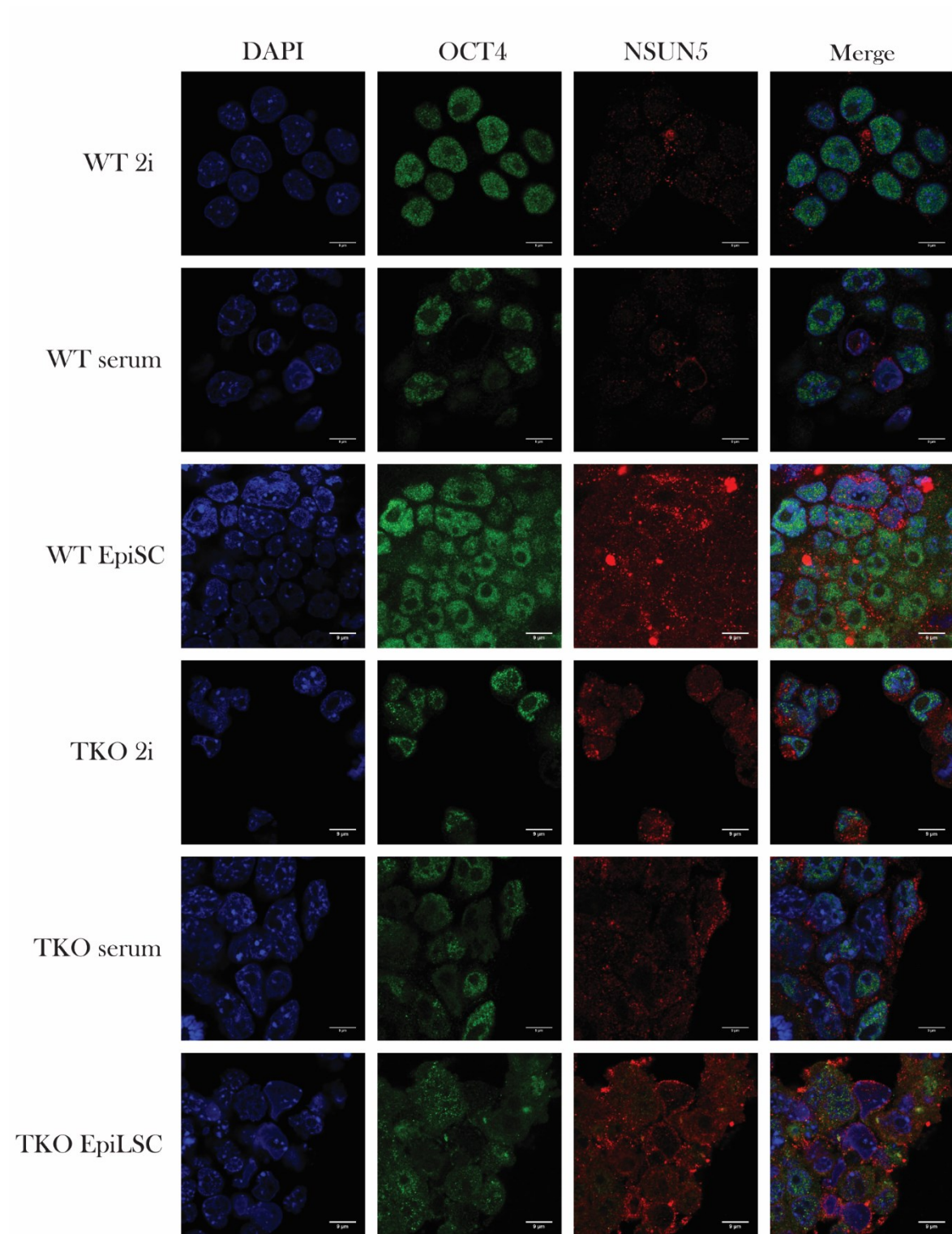


**Fig 4.11 NSUN5 subcellular localisation – WT 2i, WT serum, and TKO serum boosted signal.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu$ m.

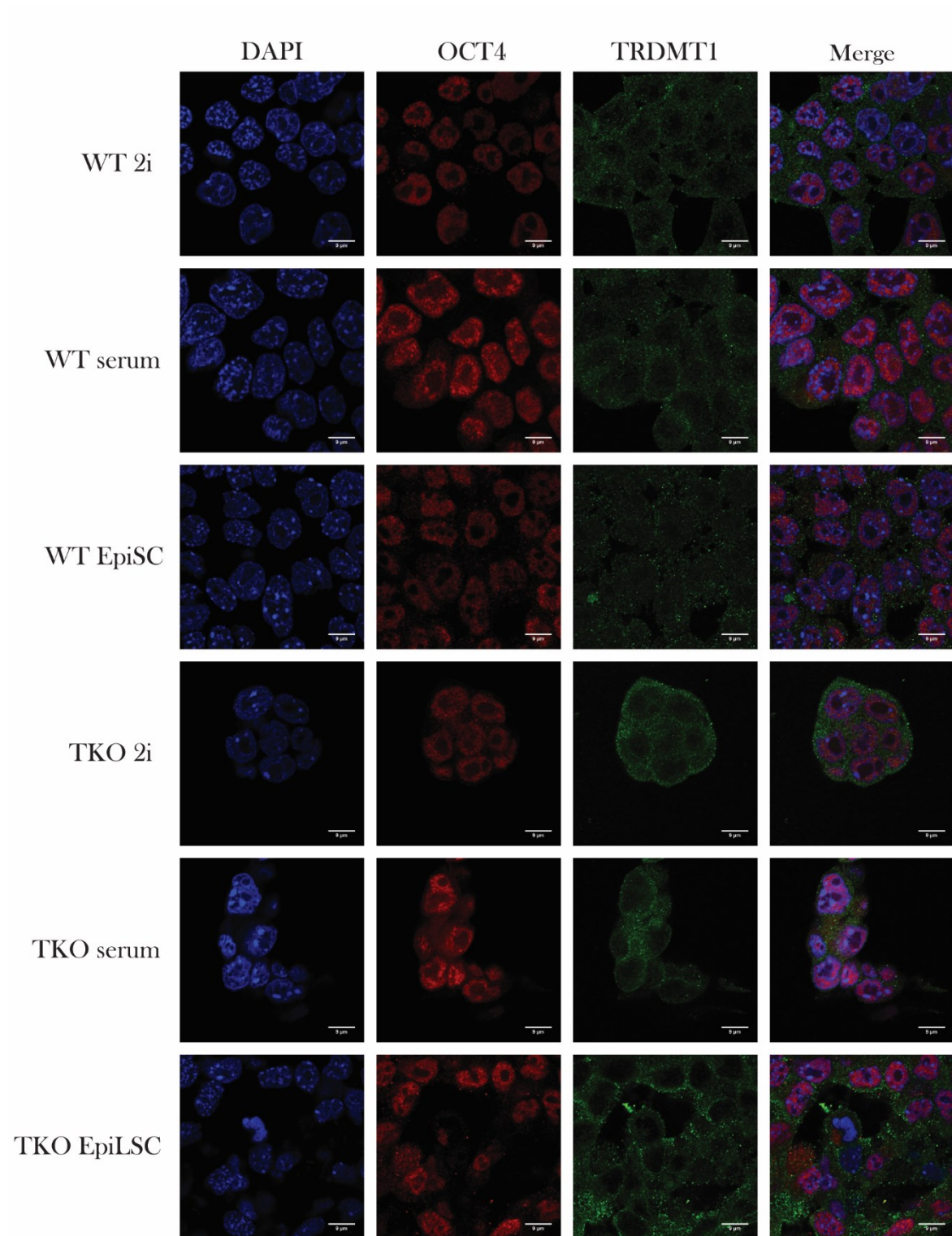


**Fig 4.12 TRDMT1 subcellular localisation.**

Stem cells were imaged using DAPI and fluorophore-coupled antibody complexes against OCT4 and TRDMT1. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of 3 biological replicates with at least 5 colonies assessed as technical replicates. J1s were used as wild type ESCs. Scale bars depicts 9  $\mu\text{m}$ .

Despite NOP2 having only ever been observed to be nucleolar, it was found to be cytoplasmic in all ESC states imaged, but not in EpiSC or TKO EpiLSC (Figure 4.7). Seemingly NOP2 is redistributed to the nucleolus from the cytoplasm during the priming of stem cells towards lineage commitment. Curiously, the TKO 2i cell state results in NOP2 existing in nuclear foci with a broader cytoplasmic signal. This captures an intermediate state between two distinct subcellular locations, but cannot provide an intermediate between ESCs and EpiSCs, as these cells are fully naïve and lack DNA methylation. This can most easily be explained by two phases of NOP2 control, in which it is nucleolar in naïve cells with very little DNA methylation, before being excluded to the nucleus during remethylation and capacitation before returning to the nucleolus upon priming. Western blotting data showed NOP2 to be most highly expressed in TKO serum cells (Figure 4.6B). The greatest immunofluorescent signal for NOP2 was observed in both wild type 2i and TKO serum cells and is a reflection of the protein being localised to small, intense foci. This highlights the advantages of quantifying protein expression by western blotting over immunofluorescent imaging of heterogeneous 3-dimensional spaces.

NSUN2 resides in the nucleolus for all cell states assayed with the exception of wild type 2i cells (Figure 4.8-4.9). As with NOP2, DNMT TKO cells grown in 2i give to an NSUN2 population split between the nucleolus and the cytoplasm. The regulation of NOP2 and NSUN2 in early naïve cells may thus be linked to give synergistic effects. That localisation to the cytoplasm is partially disrupted upon the ablation of DNA methylation is intriguing, as this is the inverse of the transition to serum and EpiSC states which bring about a completely nucleolar signal. As with NOP2, naïve hypomethylated ESCs may be transiently exporting NSUN2 to the cytoplasm before it returns to the nucleolus later in development along a naïve-primed axis. In the case of NSUN2 the return to the nucleolus occurs earlier than for NOP2, which is still cytoplasmic in serum cells. NSUN2 gives a much weaker signal in cells grown in 2i compared to serum, which

is in keeping with the western blotting data and aligns with its export to the cytoplasm being an aspect of negative regulation (Figure 4.6B).

NSUN5 is completely cytoplasmic, with no localisation dynamics observed. TKO 2i cells gave rise to the greatest signal intensity in ESCs, which is an inflection of what is observed by western blotting. Again, the concentration of epitope into foci was most likely to be the source of this discrepancy. TRDMT1 was visualised through a separate channel, as the host the primary antibody was raised in differed from that of the primary antibodies to RNMTs. To be consistent with the wavelengths used, DNMT1 images were generated using a green lookup table. TRDMT1 was exclusively observed in the cytoplasm displaying similar signal intensities in each of the ESC states (Figure 4.12). This gives rise to a static model of the expression and localisation of this enzyme which is not in keeping with western blotting data, which indicated elevated protein levels in TKO serum cells (Figure 4.6D).

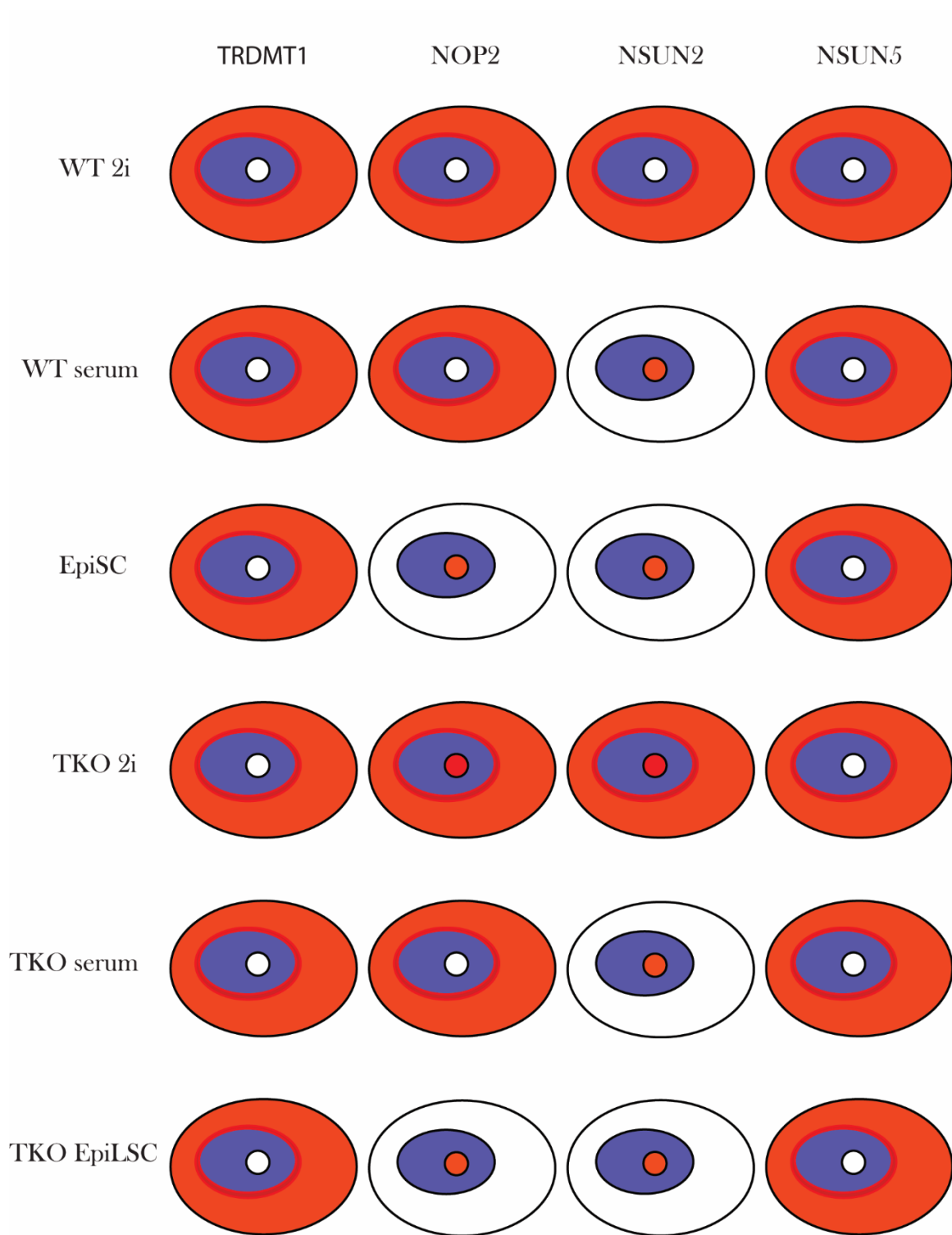


**Fig 4.13 Summation of stem cell RNMT subcellular localisations.**

The findings of immunofluorescent imaging of stem cells are depicted diagrammatically. The blue oval represents the nucleus, surrounded by the cell membrane and containing a circular nucleolus inside. The localisation of the given RNMT in the given cell state is shown in red.

### 4.3.4 RNA modification quantification

#### 4.3.4.1 Total RNA reveals distinct modification abundances along a naïve-primed developmental axis

RNA modification detection and quantification techniques, namely dot blotting and LC-MS/MS were developed and optimised as outlined in Chapter 3: Development of Methods and Tools. From that chapter, figure 3.12 shows the application of the optimised dot blotting technique for RNA 5mC detection with total RNA from the *in vitro* reprogramming ESC model. RNA was either treated with DNase, DNase and RNase, or mock-treated (without nucleases). This enabled the identification of signal due to RNA 5mC signal, in addition to any signal from contaminating DNA. Figure 3.12A gives the nature of the sample, as well as the total amount of RNA loaded for each point, while Figure 3.12B displays the chemiluminescent  $\alpha$ -5mC signal alongside the total nucleotide signal given by methylene blue staining. The DNase-treated samples all display a chemiluminescent signal which is ablated when combined with RNase treatment. These signals were thus RNA-dependent, and so considered to reflect bona fide RNA 5mC signal. Equal loading was verified by methylene blue. Imaging software was used to quantify the 5mC signal by integrating the signal intensity, with the results displayed in figure 3.12C. There is an increase in the RNA 5mC in both wild type and DNMT TKO cells grown in 2i compared to serum, while DNMT TKO does not alter the abundance of this modification (Figure 3.12A & C). This takes into account the heterogeneous background signal. These findings are consistent with RNA 5mC being regulated in line with naivety, as opposed to DNA methylation per se. While exciting on their own, LC-MS/MS was additionally used to independently verify these results.

After method development and optimisation, LC-MS/MS was performed on total RNA from wild type and DNMT TKO ESCs cultured in serum and 2i, as well as EpiSC and TKO EpiLSCs. This enabled the quantification of the abundance of cytosine, 5-methylcytosine, 5-hydroxymethylcytosine, 5-formylcytosine, adenosine and 6-methyladenosine (m6A) (Figures 4.14-16). 6-Methyladenosine has more clearly characterised roles in facilitating developmental transitions, particularly the loss of pluripotency during differentiation, than modified cytosines (Ivanova et al., 2017, Batista et al., 2014, Geula et al., 2015). For each modified nucleoside the percentage contribution to all detected forms of that nucleoside was used as a quantitative metric. This gave more biologically meaningful data than mole values standardised to input RNA amounts. Duplicate injections were performed to give technical replicates in addition to 5 biological replicates for each cell state. This enabled biologically relevant cell state comparisons, in which only genotype or culturing media relative to serum varied, to be assessed by two-way ANOVA with Tukey T-test. The integrity and ribosomal contributions of all RNA samples used was assessed using a Bioanalyzer with Nano RNA chips. A typical trace is provided in figure 4.26.

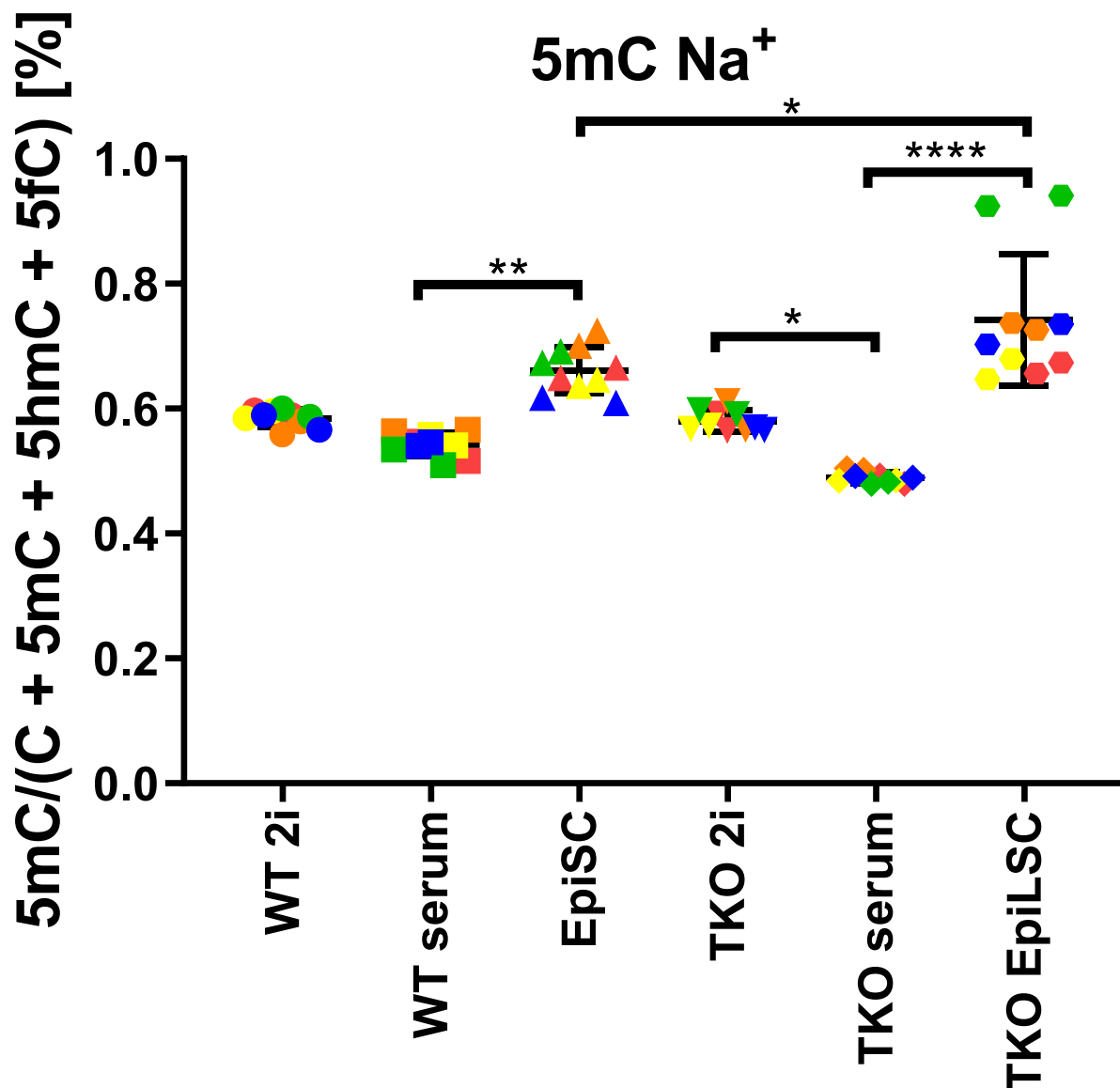


**Fig 4.14 5-Methylcytosine quantification by LC-MS/MS.**

5-Methylcytosine levels in total RNA from stem cells is provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological quintuplets (n=5) are shown, with the technical replicates matched by colour. Statistics were performed using two-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media or genotype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

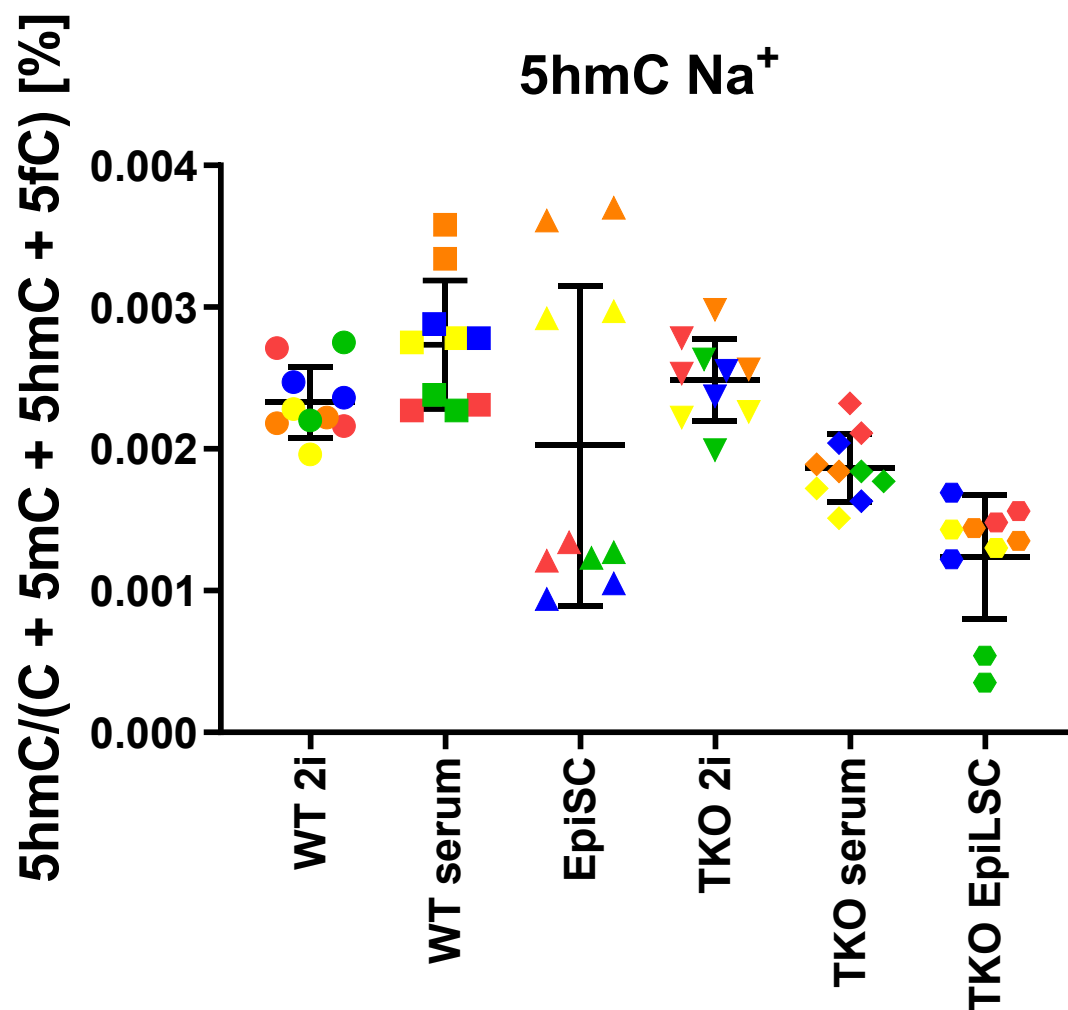


**Fig 4.15 5-Hydroxymethylcytosine quantification by LC-MS/MS.**

5-Hydroxymethylcytosine levels in total RNA from stem cells is provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological quintuplets (n=5) are shown, with the technical replicates matched by colour. Statistics were performed using two-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media or genotype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

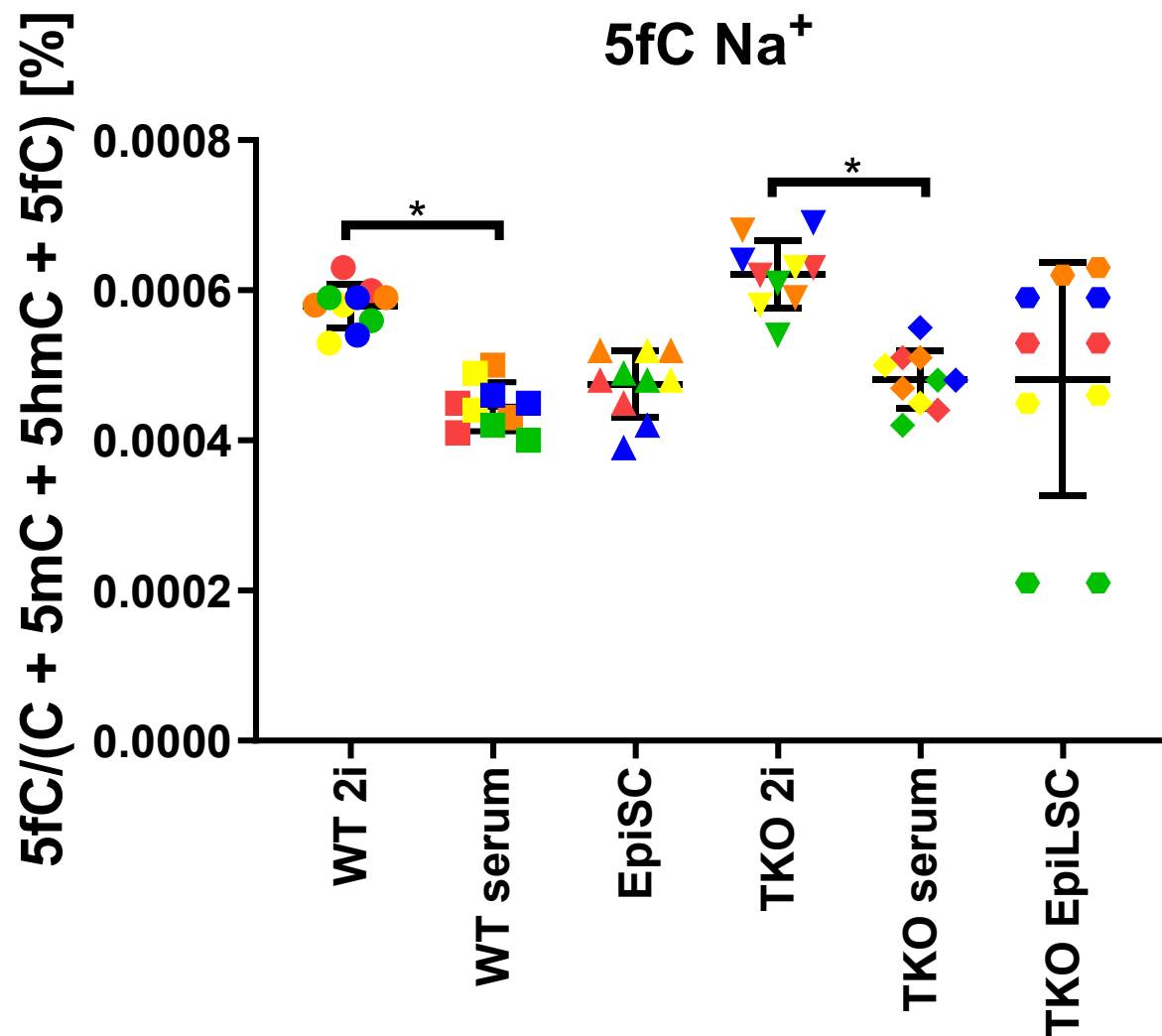


**Fig 4.16 5-Formylcytosine quantification by LC-MS/MS.**

5-Formylcytosine levels in total RNA from stem cells is provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological quintuplets (n=5) are shown, with the technical replicates matched by colour. Statistics were performed using two-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media or genotype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

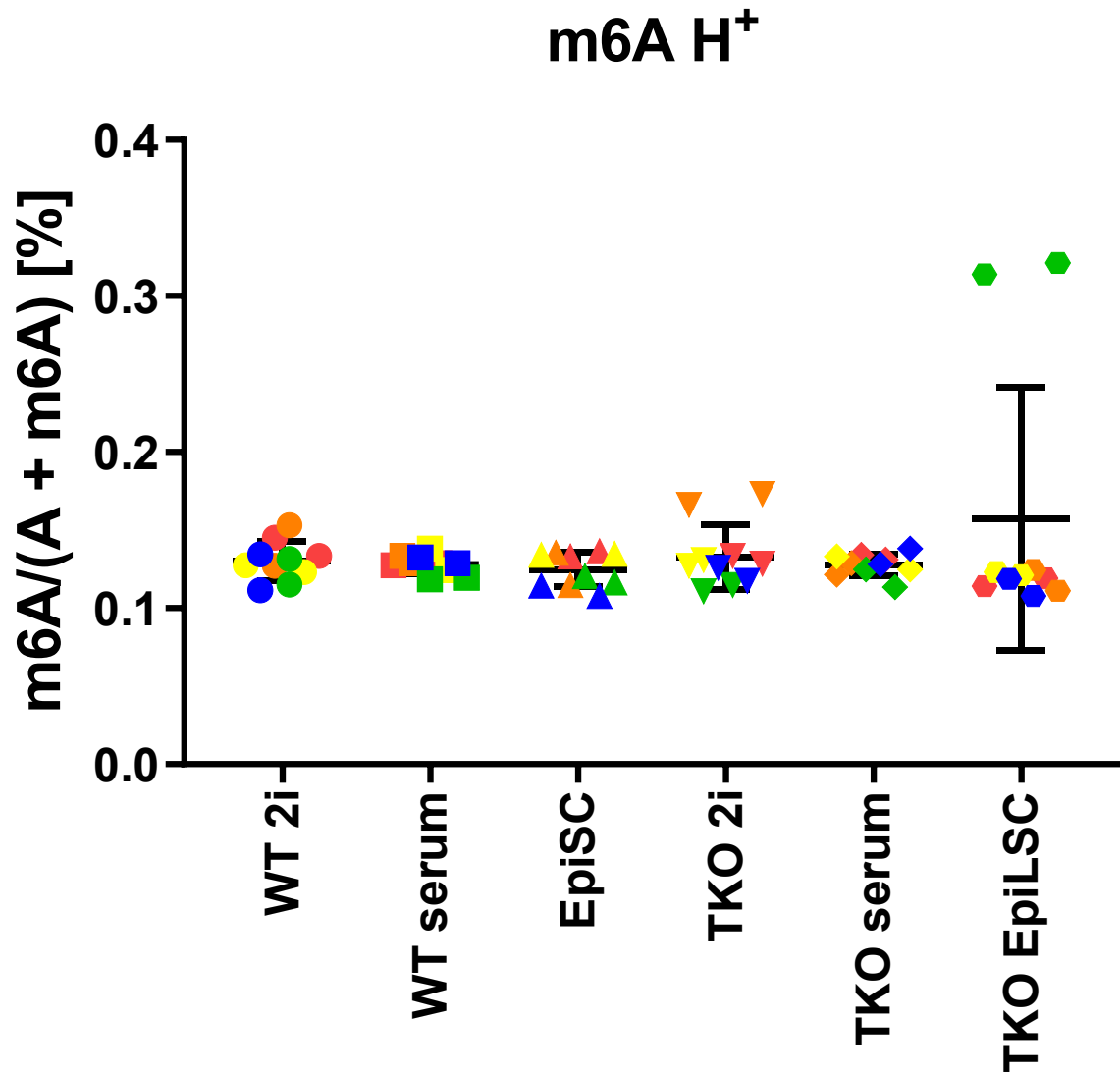


**Fig 4.17 6-Methyladenosine quantification by LC-MS/MS.**

6-Methyladenosine levels in total RNA from stem cells is provided as a percentage of all quantified forms of adenosine. J1s were used as wild type ESCs. Technical duplicates of biological quintuplets ( $n=5$ ) are shown, with the technical replicates matched by colour. Statistics were performed using two-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media or genotype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

5mC occurs at 0.6 – 0.8% of all cytosine residues in the stem cells assayed, which is slightly below what has been quantified from brain tissues (Figure 4.14) (Huber et al. 2015). The brain tissue, although heterogeneous in its cell type composition, would have mostly comprised of somatic cells. In a developmental progression from 2i, RNA 5mC levels drop in ESCs cultured in serum

before rising in EpiSC/TKO EpiLSCs (Figure 4.14). These differences are statistically significant with the exception of wild type 2i and serum grown cells. Reassuringly, higher levels of this modification in ESCs grown in 2i compared to serum is consistent with the dot blotting data displayed in Chapter 3: Development of methods and tools figure 3.12. However, this is only statistically significant in TKO cells. In addition, EpiSC and EpiLSCs display higher levels of 5mC than ESCs grown in serum. There is thus a pattern in which naïve and primed stem cells are elevated in RNA 5mC relative to a transient drop during the transitions between these states. DNMT TKO has no effect on RNA 5mC levels by either dot blot or LC-MS/MS, revealing that DNA 5mdC abundance does not intrinsically alter RNA 5mC. Rather, it is development along a naïve-primed axis which dictates RNA 5mC abundance.

5-Hydroxymethylcytosine exists at ~0.001-0.003% of all cytosine residues, while 5-formylcytosine is even less abundant at ~0.0004-0.0006% of the total cytosine residue abundance. These are incredibly low abundances, at their lowest representing a frequency of around 1 modification in 250,000 cytosines. This validates the focus on sensitivity, limit of detection, and limit of quantification during method optimisation (Chapter 3: Development of Methods and Tools) as the method was pushed to its operating limits to quantify these analytes. At these low analyte abundances the effect of noise relative to the signal is amplified, which resulted in greater variation between replicates. There were no statistically significant differences in 5hmC level, whereas 5fC was elevated in ESCs grown in 2i conditions relative to serum. This is significant in both wild type and TKO DNMT cells, which is mirrored for 5mC but not as significantly. The ratios of 5hmC/5mC and 5fC/5hmC were determined and verified that the abundance of 5hmC and 5fC was not merely a reflection of increased precursor abundance. That this is the case can be observed by eye from comparing the graphs in Figures 4.14-4.16.

6-methyladenine was quantified to levels of ~0.1-0.2%, and was notably stable in each cell state. This was surprising given the functionality of m6A in cell state and fate transitions as well as the

disintegration of pluripotency (Ivanova et al., 2017, Batista et al., 2014, Geula et al., 2015). This could be a reflection of perturbations at a RNA subtype-specific or transcript-specific level, and so small RNA and poly(A) RNA subpopulations were isolated.

#### 4.3.4.2 RNA subtypes reveal substantive and highly significant distinctions in modified RNA abundance in mRNA and small RNA

A more detailed picture of the RNA modification landscape can be revealed by delving further into the nature of RNA classes than can be achieved with total RNA samples alone. This is due to the potential for distinct RNA subpopulations to be differentially altered in their modification status. In accordance with this, RNA subtype selection was performed to enrich for poly(A) RNA and small RNA using oligo dT and cut-off columns respectively (see Materials and methods). The size cut-off for small RNAs was 200 nucleotides. Small RNA and poly(A) RNA enrichment was assessed by Bioanalyzer using Small RNA and Pico RNA chips respectively. Total RNA was also run in the Bioanalyzer to verify the integrity of all RNA samples run on the LC-MS/MS. Typical traces for these RNA preparations are given in Figure 4.26.

Modified RNAs were quantified in wild type and DNMT TKO ESCs grown in 2i and serum, as well as EpiSC and TKO EpiLSCs (Figure 4.18-4.21). In addition to this, TET triple knockout ESCs (TTKOs) were also assessed to investigate role of TET dioxygenases in RNA modification metabolism (Figure 4.22-4.25).

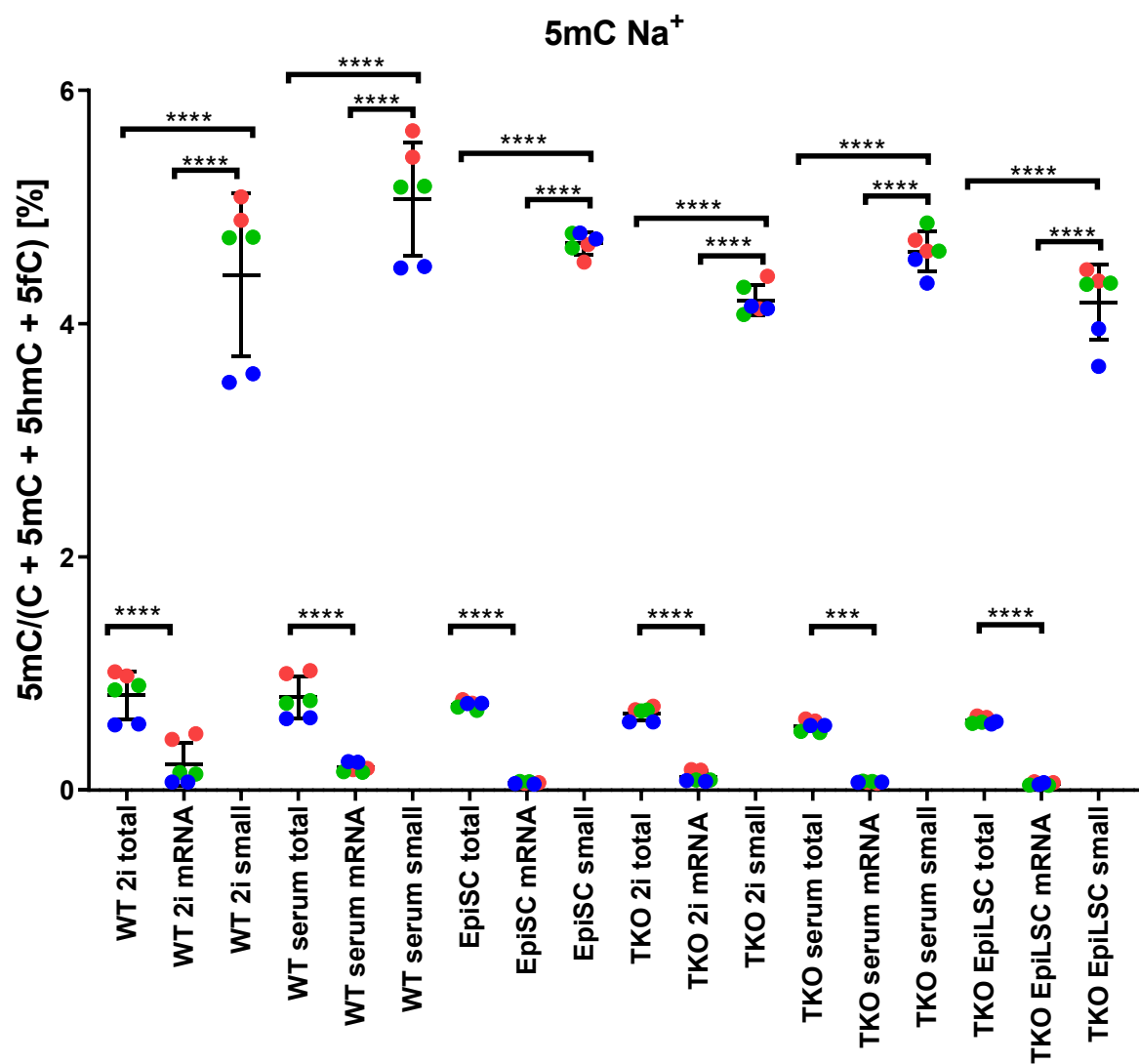


Fig 4.18 5-Methylcytosine quantification in RNA subtypes by LC-MS/MS.

5-Methylcytosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

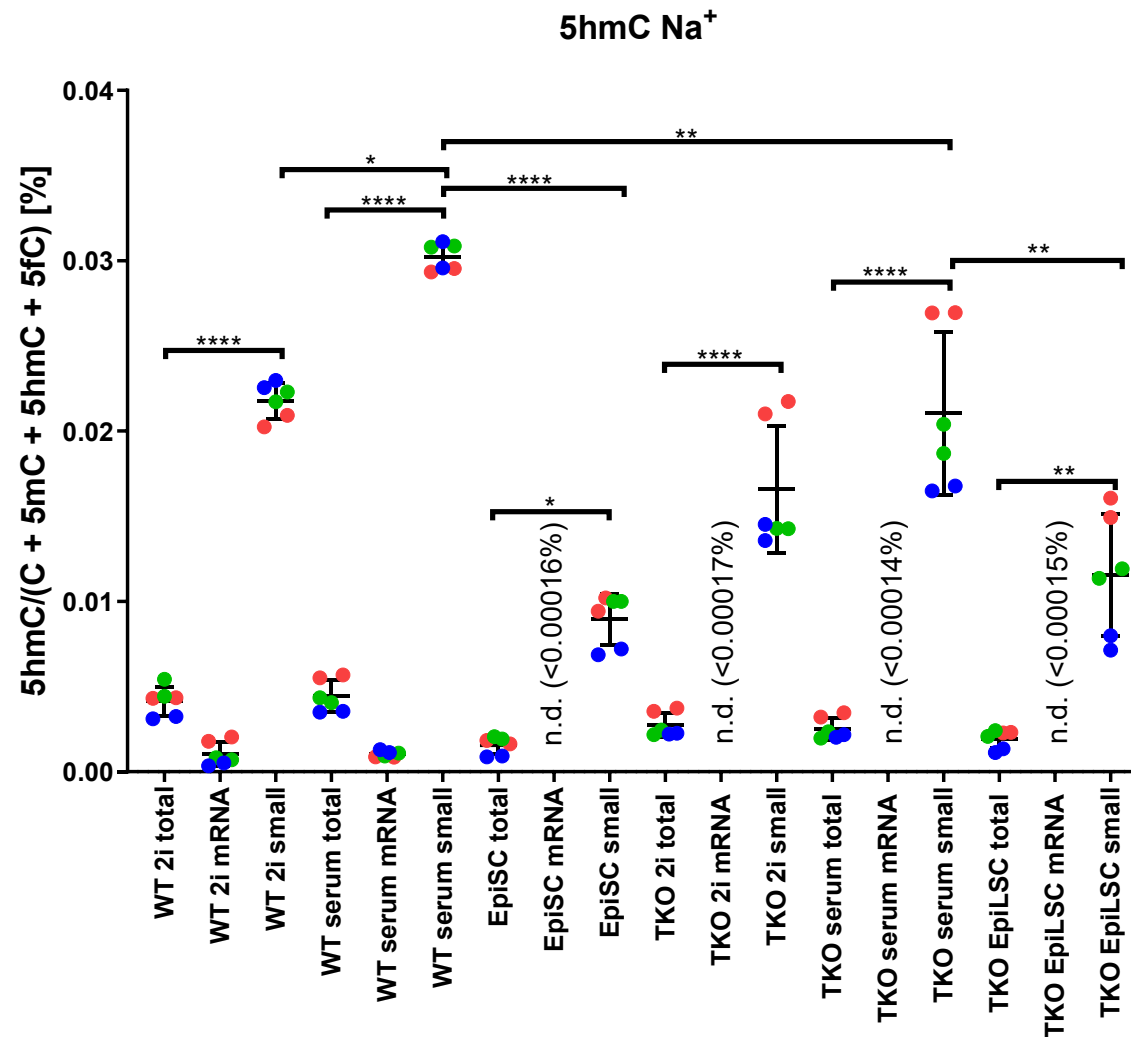


**Fig 4.19 5-Hydroxymethylcytosine quantification in RNA subtypes by LC-MS/MS.**

5-Hydroxymethylcytosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ . Where the signal fell below the limit of detection, this is indicated with n.d. (not detected) with the limit of detection given.

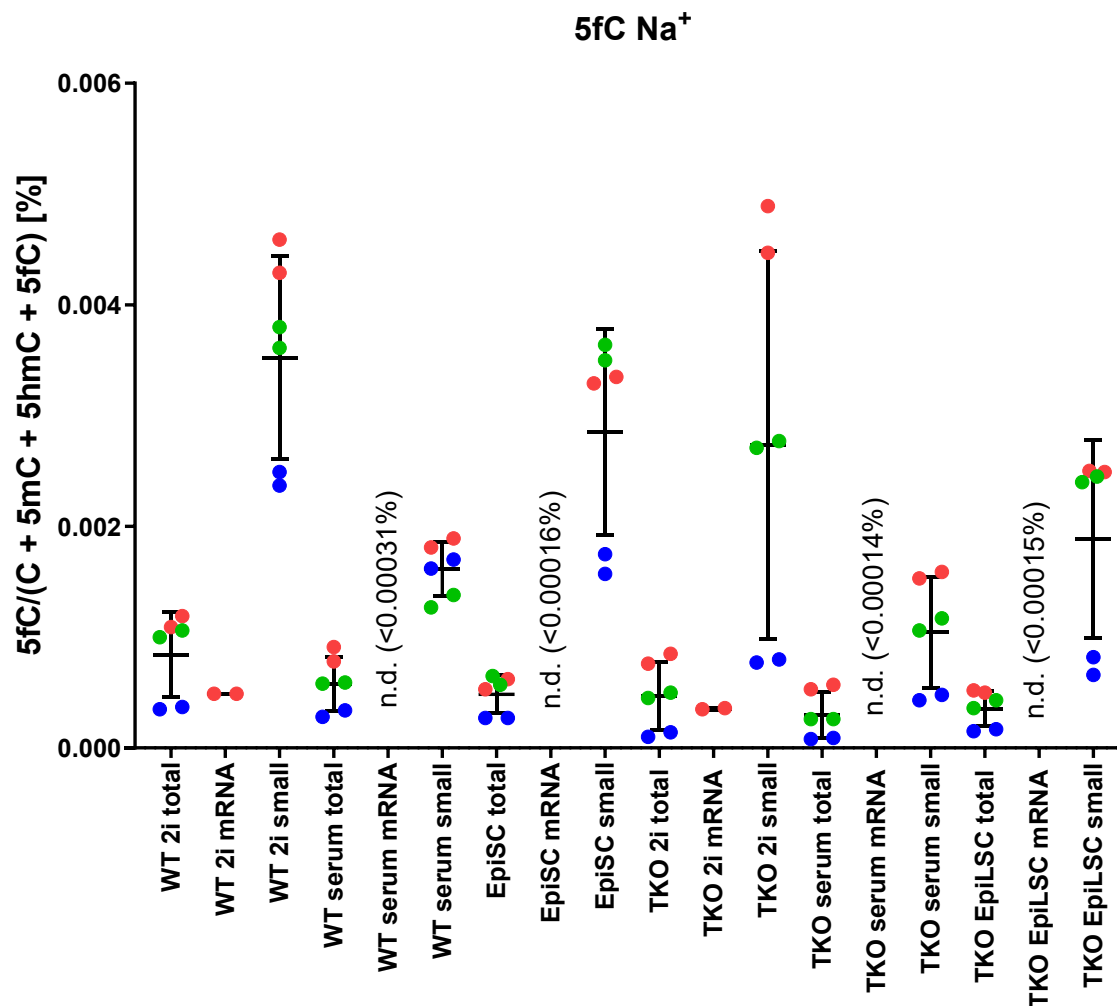


**Fig 4.20 5-Formylcytosine quantification in RNA subtypes by LC-MS/MS.**

5-Hydroxymethylcytosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ . Where the signal fell below the limit of detection, this is indicated with n.d. (not detected) with the limit of detection given.

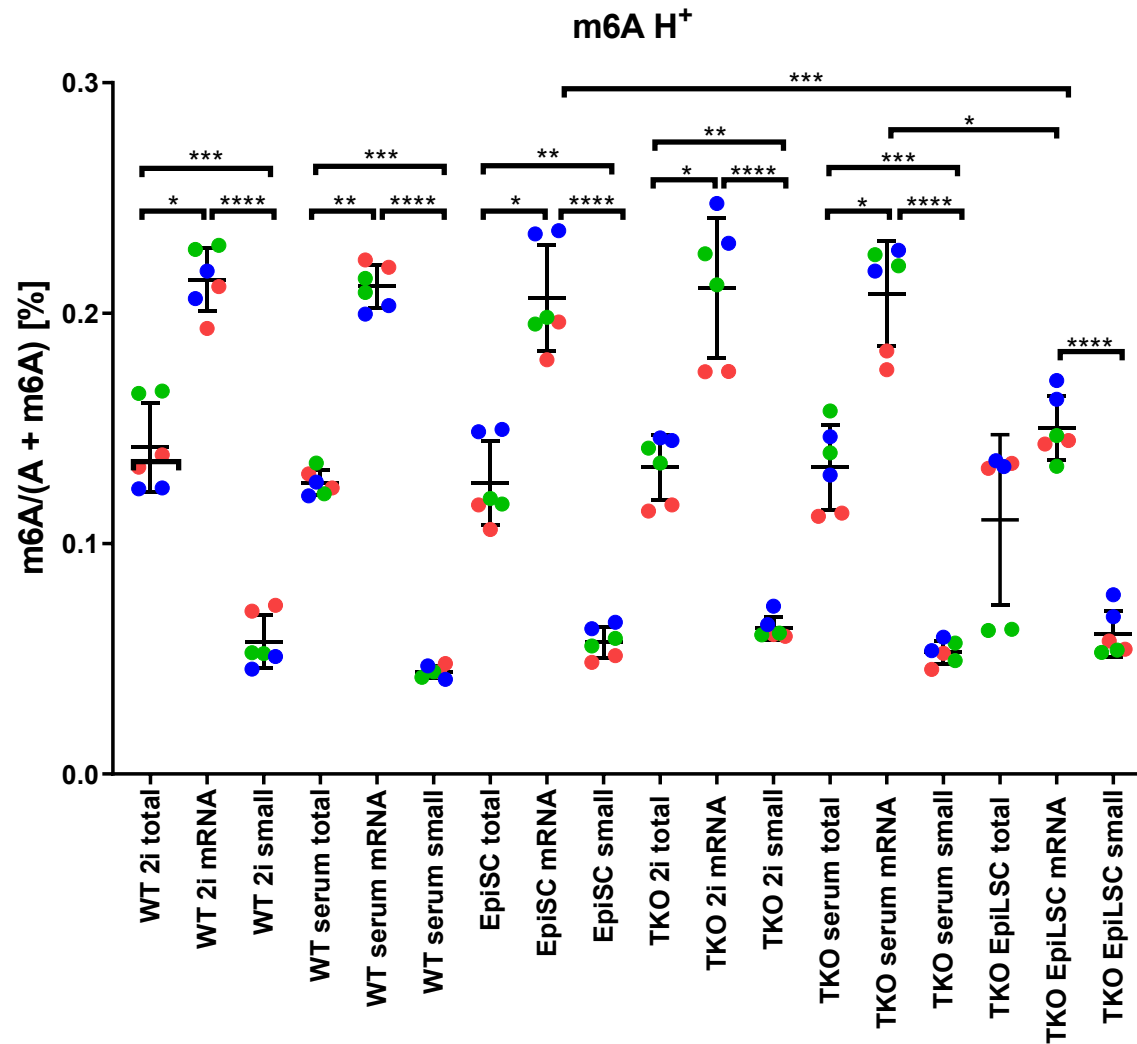


**Fig 4.21 6-Methyladenosine quantification in RNA subtypes by LC-MS/MS.**

6-Methyladenosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of adenosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

It is strikingly clear that RNA 5mC exist at disparate levels in small RNA and mRNA in relation to both the bulk average and each other (Figure 4.18). In every cell state assayed there were highly statistically significant differences for every subtype comparison. Typically, this manifested in small RNAs containing 5mC at ~5-10-fold higher levels than total RNA, which was itself ~5-10-fold higher than mRNA. In this experiment, unlike figure 4.14, there were no significant differences based upon media composition or DNMT TKO. This was most likely caused by the lower number of biological replicates (3 as opposed to 5) and an increased number of variables.

5-Hydroxymethylcytosine also shows substantive abundance differences within each cell state (Figure 4.19). Its levels in mRNA were so low as to be under the limit of detection in some samples, which highlights just how rare 5hmC is in these species, if it does exist in mRNA from these cell states at all. Only in wild type ESCs was 5hmC detectable in mRNA, and that was just at the point of limit of quantification (~0.00015%). Much like 5mC, 5hmC has around a 5-10-fold greater frequency in small RNA than total RNA. It can be observed that changes related to culturing condition or DNMT TKO are more pronounced in small RNAs, which strongly indicated that these are the target species of RNA 5hmC dynamics. The most significant difference is a drop in 5hmC in small RNAs from wild type cells transition to EpiSCs, which is also observed in TKO cells. As a result, increases in 5mC in primed EpiSC relative to serum-cultured ESCs could be explained by reduced metabolism to 5hmC. Similarly, serum-cultured ESCs display higher 5hmC levels in their small RNAs than 2i-cultured cells. This results in a comparable model for 5mC reducing through the progression from naïve ESCs cultured in 2i to a heterogeneous serum-cultured population. The only situation in which DNMT TKO has any meaningful impact on 5hmC levels is in the small RNA of serum cultured ESCs. That this is the only cell state and RNA population affected, means it was considered to be a property of small RNA from wild type serum cells displaying the highest proportion of 5hmC. There is a greater

variability in small RNA 5hmC abundances in TKO cells compared to wild types, and so it appears that the loss of DNA methyltransferases and DNA 5mdC does disrupt the normal 5hmc dynamic in the transition from naïve to primed cell state.

5-Formylcytosine followed a similar trend to 5hmC, however there were no statistically significant, biologically-relevant comparisons (Figure 4.20). This was primarily due to a combination of two factors. Abundances of 5fC existed at even lower levels in RNA than 5hmC, in samples for which both modifications could be quantified, giving an amplified signal to noise ratio relative to other nucleosides. This additionally resulted in only two single mRNA replicate samples out of the 18 analysed being quantifiable. While these samples were in each of the 2i states it would be inappropriate to imply anything from these two samples alone. When small RNAs are compared to the bulk average they are typically ~2-5-fold enriched, which is a substantial change when considered in isolation, but is more modest than what is observed for 5mC and 5hmC.

6-Methyladenosine is most abundant within mRNA and comparatively devoid in small RNA, relative to each other and total RNA (Figure 4.21). The mRNA transcript preference is consistent with published m6A-mapping data (Perry et al., 1975, Chen et al., 2015). Approximately 2-fold enrichment in mRNA compared to total RNA and ~2-fold depletion in small RNA relative to the same standard is substantive, but more modest than what is observed for modified cytosines. The data is remarkably consistent across the cell lines assayed, with the exception of TKO EpiLSCs, in which mRNA m6A levels drop. TKO EpiLSCs display elevated levels of EMT markers compared to bona fide wild type EpiSC (Figure 4.2E). As a result, some of the cells within this population progressed into differentiation, and so will have passed through the transition period of elevated m6A used to clear transcripts and facilitate the adoption of a new cell state.

Together, these results are telling. They show that the primary factor dictating modified RNA abundance is the RNA subtype, with modified cytosines being most prevalent in small RNA and m6A is more concentrated within mRNA transcripts. Additionally, 5hmC transiently increases in the transitions between naïve and primed cells in a manner which is partially disrupted by the absence of DNA methyltransferases. Reduced m6A levels in TKO EpiSCs are a reflection of their partial differentiation. Further experimentation was performed in TET triple knockout cells with the expressed purpose of identifying the role of these enzymes in modified cytosine metabolism.



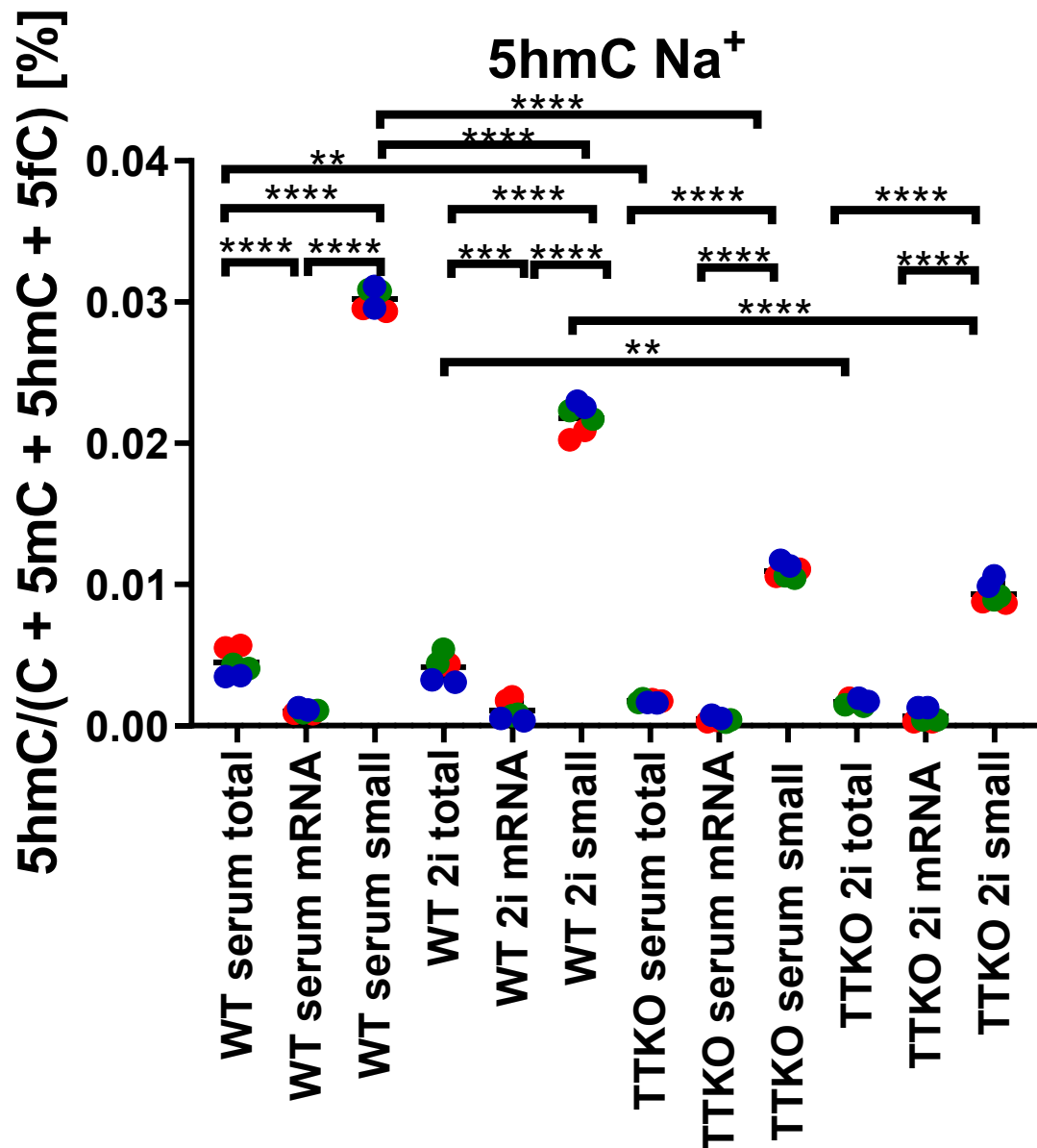


Fig 4.23 5-Hydroxymethylcytosine quantification by LC-MS/MS in TET triple knockout ESCs.

5-Hydroxymethylcytosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

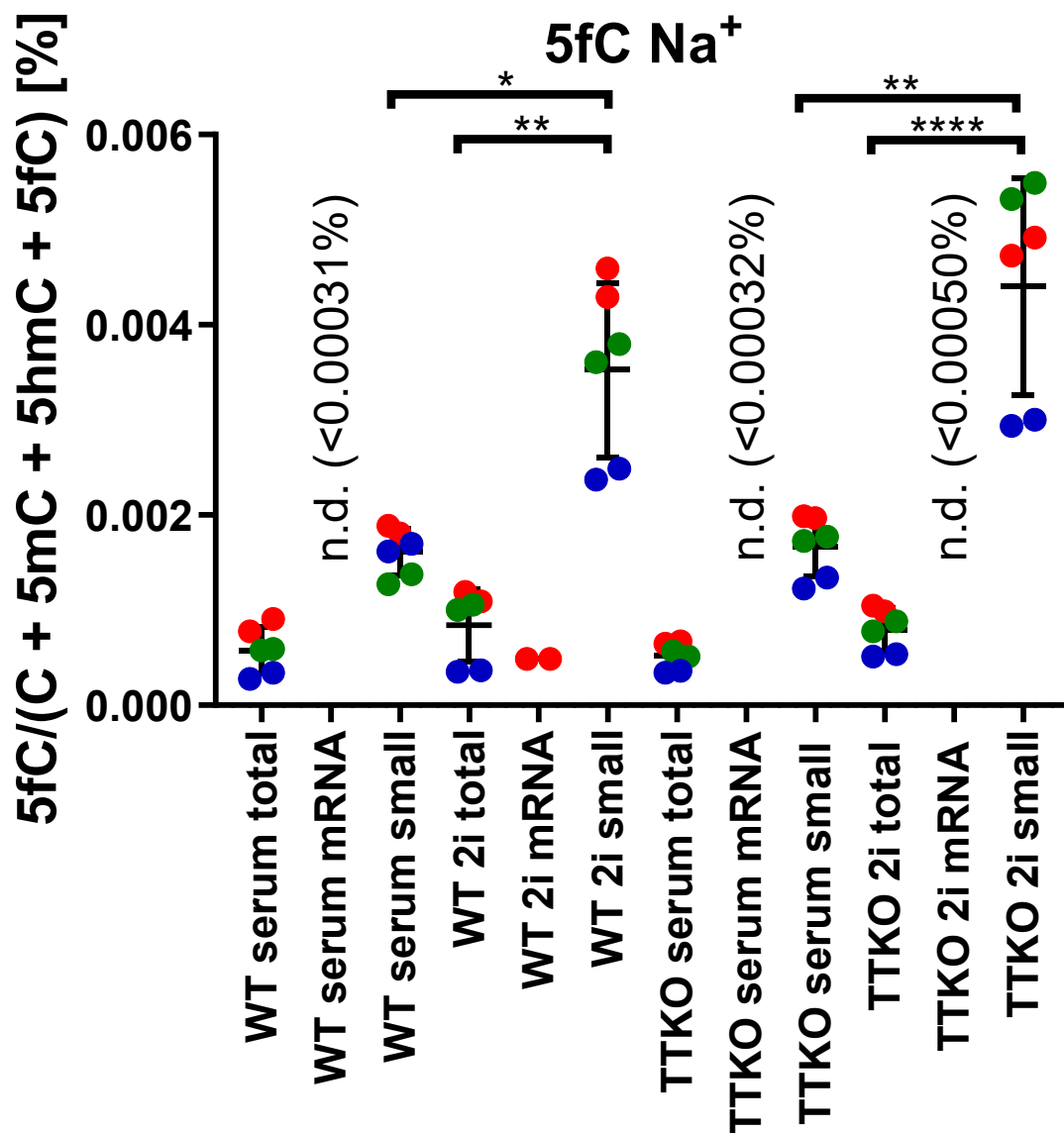


**Fig 4.24 5-Formylcytosine quantification by LC-MS/MS in TET triple knockout ESCs.**

5-Formylcytosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of cytosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ . Where the signal fell below the limit of detection, this is indicated with n.d. (not detected) with the limit of detection given.

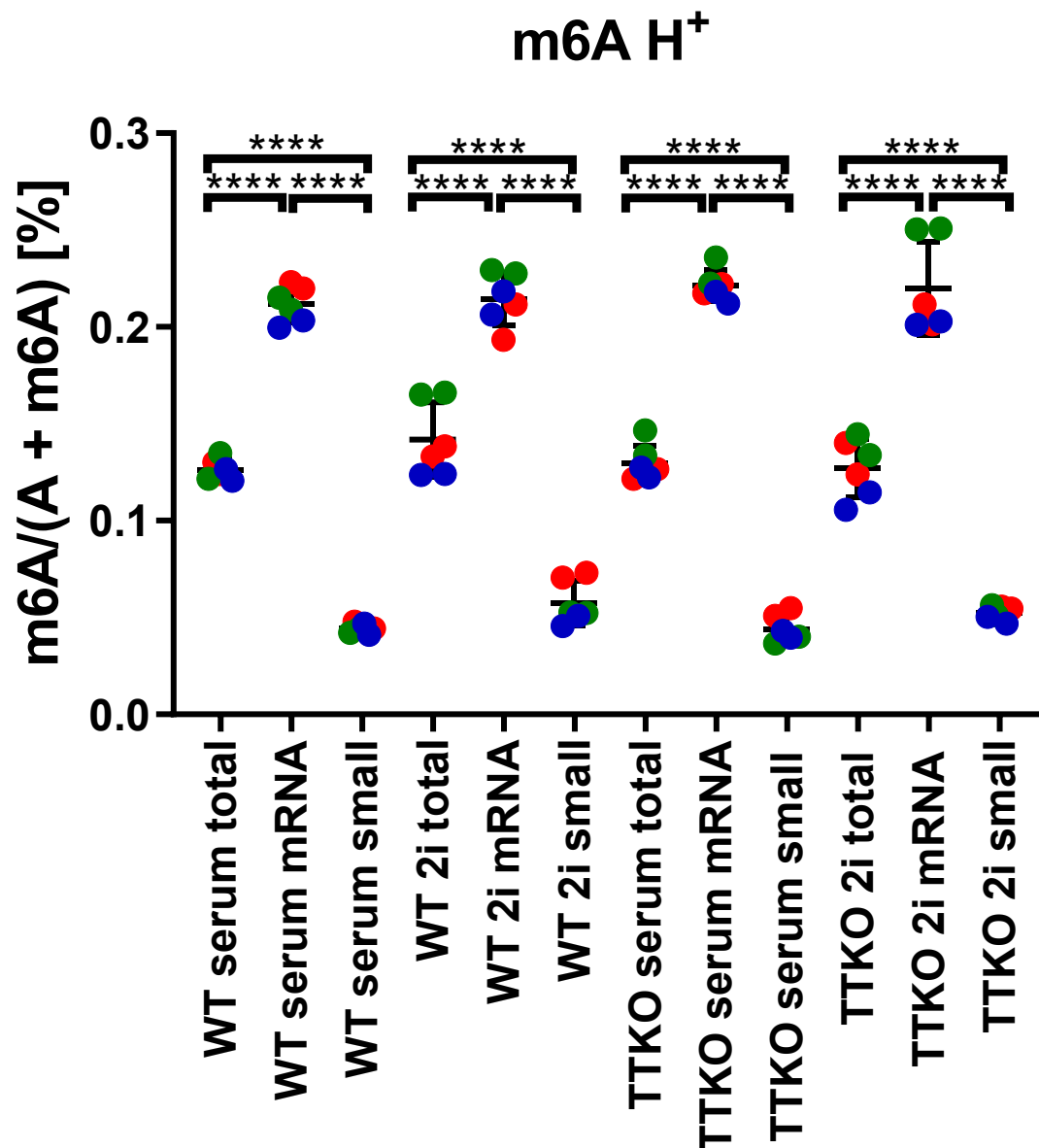


**Fig 4.25 6-Methyladenosine quantification by LC-MS/MS in TET triple knockout ESCs.**

6-Methyladenosine levels in total RNA from stem cells are provided as a percentage of all quantified forms of adenosine. J1s were used as wild type ESCs. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. Statistics were performed using three-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only culturing media, genotype, or RNA subtype varied. Where statistically significant differences occur, they are demarcated as follows: \* =  $p < 0.05$ , \*\* =  $p < 0.005$ , \*\*\* =  $p < 0.0005$ , \*\*\*\* =  $p < 0.0001$ .

TET triple knockout had no effect on RNA 5mC levels, while the same distinctions based upon RNA subtype observed in Figure 4.17 were found, save for mRNA compared to total RNA (Figure 4.22). This was no longer statistically significant in wild type cells. Upon inspection of the formulae used, the difference arose from higher confidence intervals for the difference between the means. This was due to an overall smaller and more variable total dataset. Regardless, the same pattern of lower 5mC levels in mRNA compared to total RNA can still be observed.

TTKO cells have lower 5hmC levels in both total and small RNA in serum and 2i media conditions (Figure 4.23). This shows the TETs to be involved in RNA 5hmC biosynthesis, but critically indicates an independent synthesis route. Such a route must exist, as RNA 5hmC exists in the complete absence of functioning TET enzymes. 5-Hydroxymethylcytosine was quantifiable in mRNA in the absence of TET enzymes, but not in the absence of DNMTs or in EpiSC. When juxtaposed in this way, the regulation of 5hmC along a naïve-primed developmental axis in a DNMT-dependent manner becomes more striking.

The abundance of 5fC was unaltered by TTKO, and so an alternative catalytic mechanism must cause its formation within ESCs (Figure 4.24). 5-Formylcytosine is elevated in small RNA in 2i conditions compared to either total RNA or small RNAs from serum cells. This was observed previously, but holds greater statistical weight in this dataset (Figure 4.20).

The knockout of the TET enzymes has no effect on m6A levels, which was as expected (Figure 4.25). The differences observed previously in mRNA, total, and small RNAs in ESCs were also observed in this experiment (Figure 4.21).

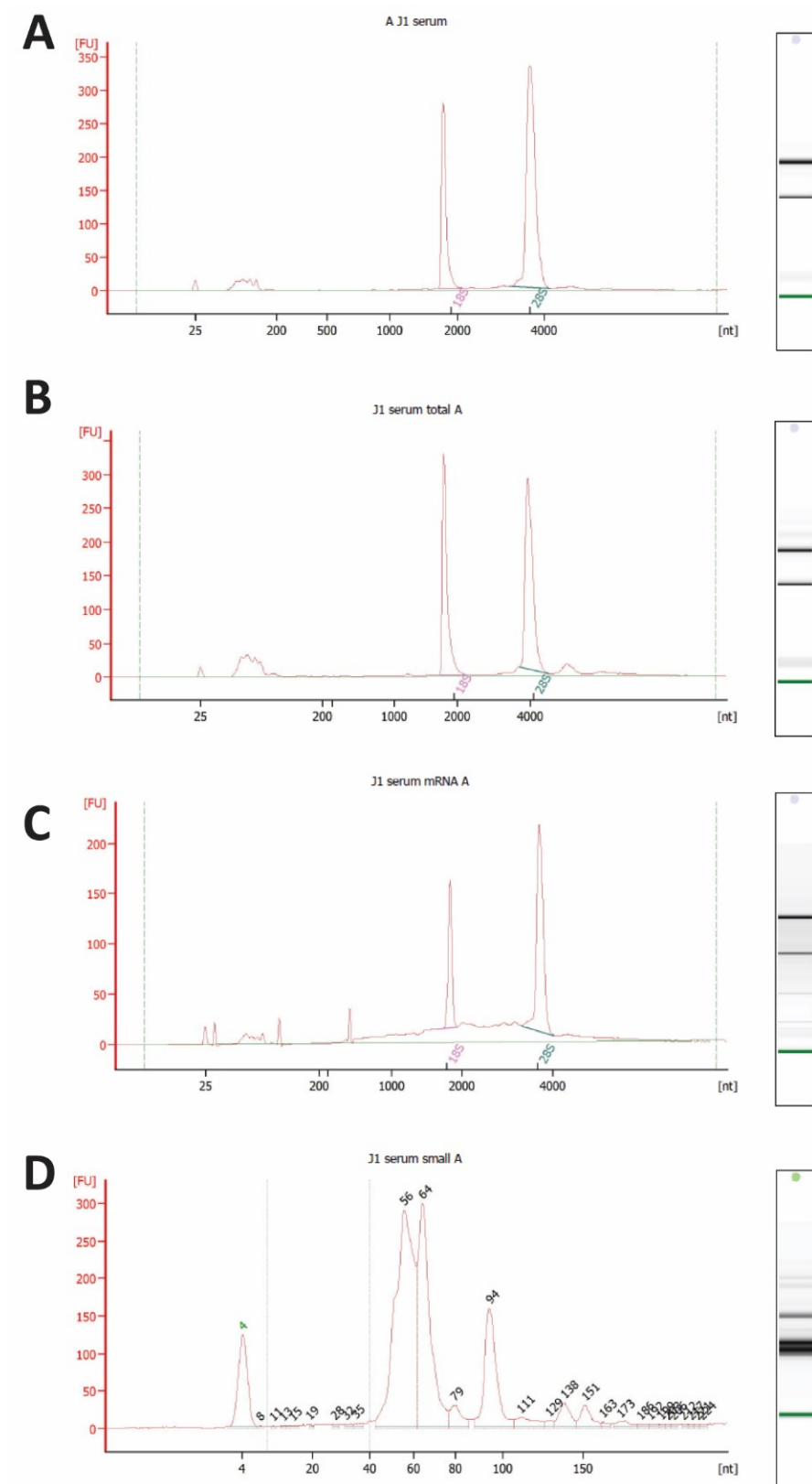


**Fig 4.26 Representative Bioanalyzer traces.**

Representative Bioanalyzer traces are given for total RNA (**A**), total input RNA (**B**), mRNA (**C**) and small RNA (**D**). In each case, RNA from ESCs grown in serum are provided.

## 4.4 Discussion

Experimentation in stem cells was conducted to investigate how modified RNAs, particularly 5-methylcytosine and its oxidised derivatives, and their deposition machinery might be related to DNA methylation status. This was disentangled from, and compared to, the effects of adopting a naïve cell fate, as occurs during epigenetic reprogramming. Additionally, the processes involved in capacitation and priming, modelled by ESCs grown in 2i, ESCs grown in serum, and EpiSCs, were also used to model RNA modification and RNMT dynamics. A final stem cell system, in which all three TET enzymes were ablated, facilitated the investigation of these enzymes in the oxidation of 5mC.

In the absence of DNA methylation, mammalian cells cannot stably adopt differentiated cell states (Sheaffer et al., 2014). Considering this, it is in some ways surprising that in DNMT TKO cells an EpiSC-like state can be induced. Priming may thus be the final boundary to which mammalian cells lacking DNA methylation can reach before becoming destabilised and undergoing apoptosis. That some elevation of EMT markers is observed in TKO EpiSCs provides a tentative clue that these cell may be able to enter the differentiation process in some way, but this is not canonical. Recapitulation of an EpiSC-like state lacking DNA methylation was novel to this work.

The RNMTs NOP2, NSUN5, and TRDMT1 are upregulated in the absence of DNA methylation in a manner that is ablated by the naivety of 2i culture. Such naivety is associated with lower levels of NSUN5 and TRDMT1 expression. From this it was hypothesised that RNA 5mC would increase in abundance upon the loss of DNA methylation, while being reduced in naïve cells through a mechanism independent of DNA methylation. However, the converse is true. By both dot blotting and LC-MS/MS, it was found that 5mC is independent of DNA methylation and increased in naïve cells cultured in 2i. Regulation of RNA modification level is

therefore not brought about by modulation of the abundance of these RNMTs. The subcellular localisation of these enzymes reveals unexpected cytoplasmic localisations for NOP2, NSUN2, and NSUN5. NOP2 returns to the nucleolus in primed EpiSC and TKO EpiLSCs. This is consistent with a role in bringing about the increases in 5mC observed in these cells, through increased catalytic turnover in the nucleolus.

The key determinant of the RNA 5mC, 5hmC, 5fC and m6A level is the RNA subtype analysed in each cell state assayed. This highlights the functional and regulatory divergence of these differing RNA classes, with the modified cytosines primarily located in small RNAs and m6A in mRNA. RNA 5hmC and 5fC exist at very low abundances, particularly in mRNA. Many of the samples analysed were below the limit of quantification, even for the highly sensitive LC-MS/MS methodology optimised for this approach. A transient increase in RNA 5hmC in the heterogeneous serum-cultured cell state between naïve and primed cells results in a modest reduction in its precursor 5mC at this stage. While these developmental changes were both novel and exciting, rather perplexingly m6A was highly static. 6-methyladenosine was only altered within any given RNA pool in mRNA from TKO EpiLSCs. The reductions observed in these cells are indicative of part of the TKO EpiLSCs population undergoing differentiation, and so having already cleared the transcripts of the previous pluripotent state.

DNMT triple knockout almost universally did not alter modified cytosine levels. The sole exception was 5hmC in the small RNA of serum cultured cells, which in wild type cells represents the greatest concentration of 5hmC measured. When considered together, this is a modest effect for something as fundamental as the complete removal of DNA methylation. TET triple knockout did not alter 5mC or 5fC abundances, while 5hmC levels were reduced, particularly in small RNAs, but not ablated. These results necessitate a separate, yet to be identified enzyme(s) in the formation of 5hmC and 5fC.

It is clear that DNA 5mdC and RNA 5mC are not intrinsically interrelated in the manner first hypothesised. Rather, 5mC, 5hmC, and 5fC are regulated in an RNA subtype-specific manner by the developmental progressions that occur between naïve pluripotent cells and primed ones. This is accompanied by the RNMT NOP2 localising to the nucleolus where it could mechanistically drive these changes. Equipped with this knowledge, *in vivo* work was performed to more fully understand these processes within a truly biologically relevant context. This is outlined in the next chapter.

# Chapter 5: RNA modifications

## *in vivo*

## 5.1 Introduction

A murine model was chosen to investigate RNA modifications within an *in vivo* DNA demethylation context brought about by epigenetic reprogramming. Within the mammalian life cycle there are three instances of this occurring, which can be broken down into either zygotic/early embryo, or germline reprogramming (Introduction and literature review: Figure 1.12) (Saitou et al., 2011). In the zygote the male pronucleus loses DNA 5mdC immediately post-fertilisation, while the maternally inherited DNA loses its methylation more slowly during early embryo development from the 1-cell stage to the morula or early blastocyst stage (Mayer et al., 2000, Amouroux et al., 2016, Kafri et al., 1992, Monk et al., 1987, Howlett et al., 1991). The germ line is specified in the proximal epiblast before residing in the hindgut, and then migrating to the developing gonad (Lawson et al., 1999, Anderson et al., 2000, Tam et al., 1992, et al., 2002). Upon entry to this tissue, germ cells undergo epigenetic reprogramming with the concurrent reductions in DNA methylation (Hajkova et al., 2002, Hill et al., 2018). The duration over which active demethylation occurs is typically 24 hours, however the entry of PGCs to the genital ridge is not synchronised (Hajkova et al., 2002, Hill et al., 2018). Accordingly, DNA demethylation occurs between E10.5 and E12.5. In none of these processes is DNA methylation completely ablated, as is observed in DNMT triple knockout cells. Within these systems the cell numbers, and resultant amounts of material, provided constraints upon what experimentation was feasible.

Primordial germ cells were chosen as the main *in vivo* model, as they provide the greatest cell number out of the 3 *in vivo* contexts available. PE OCT4GFP males were crossed with wild type females, of C57bl6 background, to give heterozygous embryos with GFP-labelled PGCs in accordance with the restricted expression of OCT4 to these cells (Yoshimizu et al., 1999). This

facilitated the identification of PGCs from surrounding somatic tissue as well as the isolation of pure populations by fluorescence-activated cell sorting (FACS) (Hajkova et al., 2008).

## 5.2 Experimental Aims

In order to initially characterise the systems, I sought to determine the expression of key genes relating to RNA 5mC and m6A in *in vivo* reprogramming. Subcellular localisation was of key interest following on from *in vitro* work (Chapter 4: RNA Methylation in Stem Cells). As such, it was deemed a priority to establish where NOP2, NSUN2, NSUN5 and TRDMT1 localise within cells undergoing *in vivo* reprogramming. In doing so, this enabled a resolution as to whether the subcellular localisation changes observed *in vitro* reprogramming recapitulated *in vivo*. Quantification of RNA modification levels in *in vivo* reprogramming was considered a priority in understanding their function within this context.

## 5.3 Results

### 5.3.1 Gene expression analysis reveals dynamic changes in RNMT and m6A-related gene expression in early embryo and PGC development.

RNA-seq datasets previously generated in the lab were used to ascertain the expression of RNMTs and m6A readers, writers, and erasers. These datasets covered epigenetic reprogramming in the zygotic maternal pronucleus during early embryo development, and primordial germ cell epigenetic reprogramming (B. Nashun, Hill et al, 2018).

In PGCs undergoing epigenetic reprogramming *Nop2* is highly expressed at a level above the other RNMTs (Figure 5.1A). With reads per kilobase million (RPKM) values above 60, this falls into one of the highest percentiles for gene expression within these cells (P. Hill, personal communication). When the other RNMTs are considered each is found to be expressed, some

at moderately high levels of 10-20 RPKM (Figure 5.1B). *Nop2*, *Trdm1*, *Nsun2*, and *Nsun5* and all display positive trends around the time course of epigenetic reprogramming. These are the same RNMTs that were found to be upregulated in the *in vitro* reprogramming model, and which were analysed in stem cells (Chapter 4: RNA methylation in stem cells). While this was encouraging, the m6A readers, writer, and erasers showed little variation over the developmental time course. The notable exception to this is the incremental increase in the reader protein *Ythdf2*. *Ythdc1* expression was not detected in any of the samples. YTHDC1 is involved in SAM homeostasis through degradation of the *Mat2a* gene encoding the SAM synthase (Mendel et al., 2018, Shima et al., 2017). Consequently, as YTHDC1 is absent it cannot be involved in any SAM homeostasis during this time of rapidly changing methylation landscape.

In the early embryo dataset there were sufficient replicates to enable statistical comparisons (Figure 5.1D-G). *Nop2* expression increases ~100-fold from the 1-cell to 2-cell stage, before progressively dropping off. *Trdm1* and *Nsun4* are expressed most highly in the 1-cell stage, before experiencing large proportional decreases in their mRNA abundance. In the instance of *Trdm1*, a bimodal pattern emerges, in which the mRNA abundance is transiently increased during the morula stage. *Nsun2* undergoes a gradual transient wave of upregulation, reaching a maximum in the 4-cell stage. It is clear that there is a highly dynamic RNMT landscape during the early embryo development. This was considered encouraging in relation to the roles these enzymes may play during maternal DNA demethylation and early embryo development more broadly. The picture of m6A readers, writers, and erasers in early embryo reprogramming is not as clear as that for RNMTs (Figure 5.1F-G). However, there is a notable progressive reduction in the expression of the methyltransferase *Mettl3* and reader protein *Wtap* at the mRNA level. *Ythdf2* undergoes a wave of upregulation which peaks at the 4 cell stage. This protein is necessary for the maternal to zygotic transition that enables development beyond the 2-cell stage (Ivanova et al., 2017).

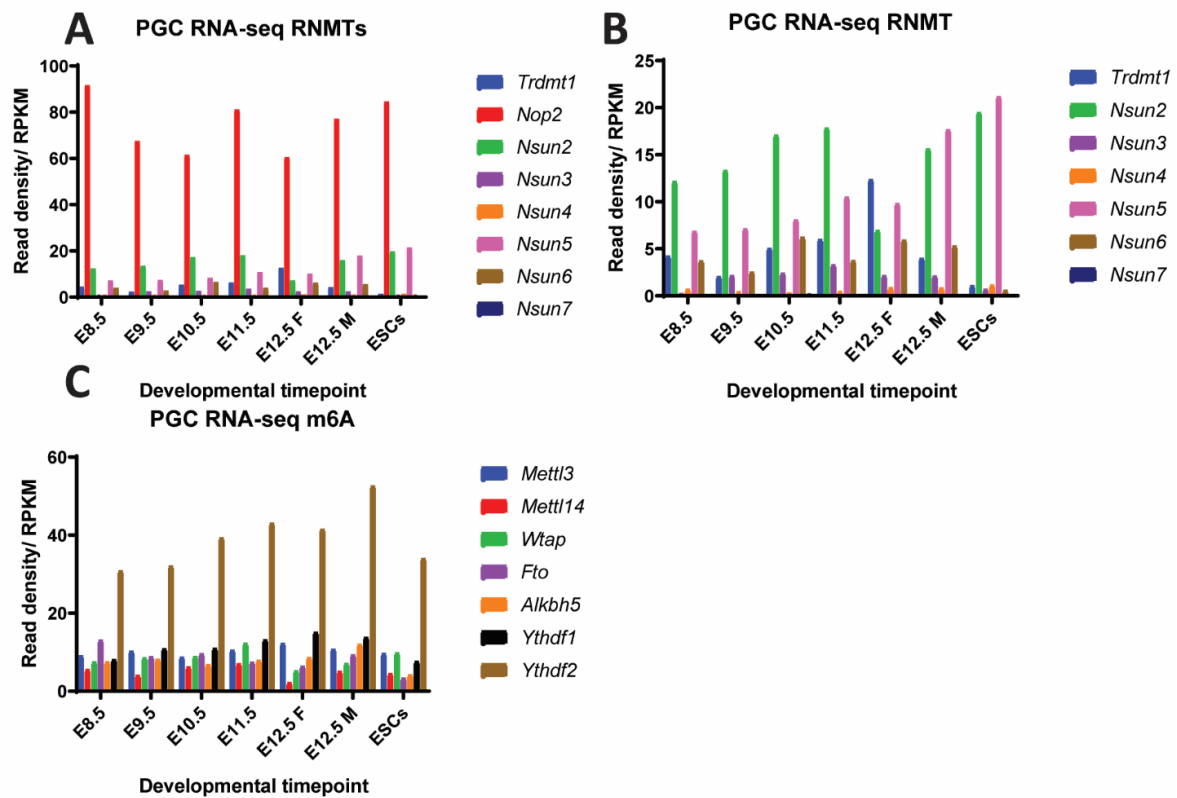


**Fig 5.1 PGC RNA-seq reveals dynamic expression changes of proteins involved in 5mC and m6A metabolism during epigenetic reprogramming.**

(A-C) Previously published PGC RNA-seq data are presented (Hill et al, 2018). The detected RNMTs (A-B) and m6A readers, writers, and erasers (C) are displayed. The data in (A) is reproduced in (B) without *Nop2* due to the large range of count values between the RNMTs. The data presented is representative of n=2 biological replicates.

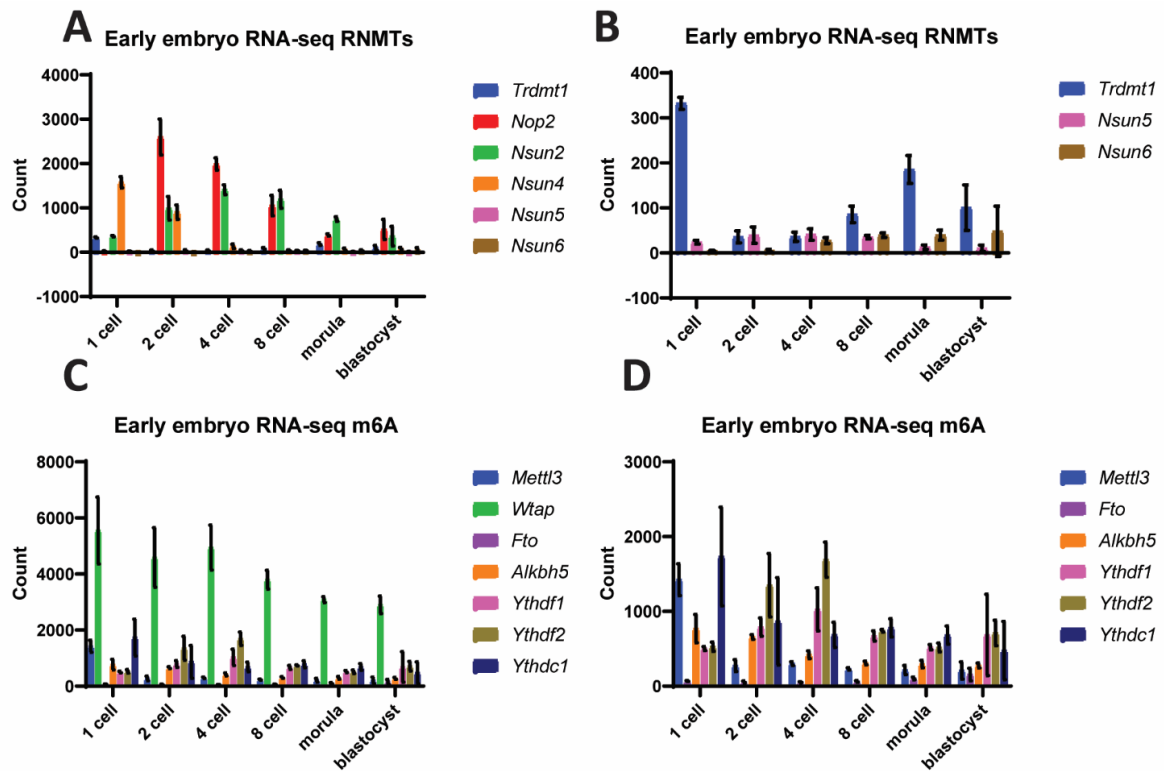


**Fig 5.2 Early embryo RNA-seq reveals dynamic expression changes of proteins involved in 5mC and m6A metabolism during epigenetic reprogramming.**

**(A-D)** Early embryo RNA-seq data generated by a colleague are presented (B. Nashun). The detected RNMTs **(A-B)** and m6A readers, writers, and erasers **(C-D)** are displayed. The data in **(A)** is reproduced in **(B)** without *Nop2*, *Nsun2*, and *Nsun4* due to the large range of count values between the RNMTs. The data in **(C)** is reproduced in **(D)** without *Wtap* due to the large range of count values between the m6A readers, writers, and erasers. Error bars display standard deviation calculated from n=3 replicates.

### 5.3.2 RNMT protein levels and subcellular localisations in early embryo and PGC development

#### 5.3.2.1 Primordial germ cells have elevated RNMT protein levels and display dynamic changes in NOP2, NSUN2, and NSUN5 subcellular localisations

At the nadir of DNA demethylation in E12.5 PGCs, the total germ cell population is approximately 1000 cells per embryo. This decreases precipitously in earlier developmental time points. As such, substantial pooling of samples would be necessary to perform western blotting. While recent advances have brought such experimentation into the realms of feasibility, reproducibility is a problem for low input western blotting (Dr. Tien-Chi Huang, personal communication). In light of this, immunofluorescent staining was used to quantify protein abundance and subcellular localisation.

Initial experimentation found that the  $\alpha$ -RNMT antibodies available performed poorly on genital ridge cryosections, particularly in regards to permeabilisation into the tissue. This did not resolve upon optimisation of a range of relevant parameters, and so single cell genital ridge immunofluorescent staining was performed. Having previously established protocols for immunofluorescent imaging of stem cell colonies using the  $\alpha$ -RNMT antibodies, resorting to a single-cell approach enabled imaging. By the nature of dissociating cells from dissected genital ridges without cryopreservation, it was not possible to include multiple embryonic time points on the same slides. As a result, the expression of the RNMTs in each panel cannot be compared to other time points, but expression relative to the surrounding somatic cells can be observed. This data is displayed in figures 5.2-5.9 at a lower magnification than immunofluorescence images in figures 5.10-5.16 to encompass more cells within the field of view.

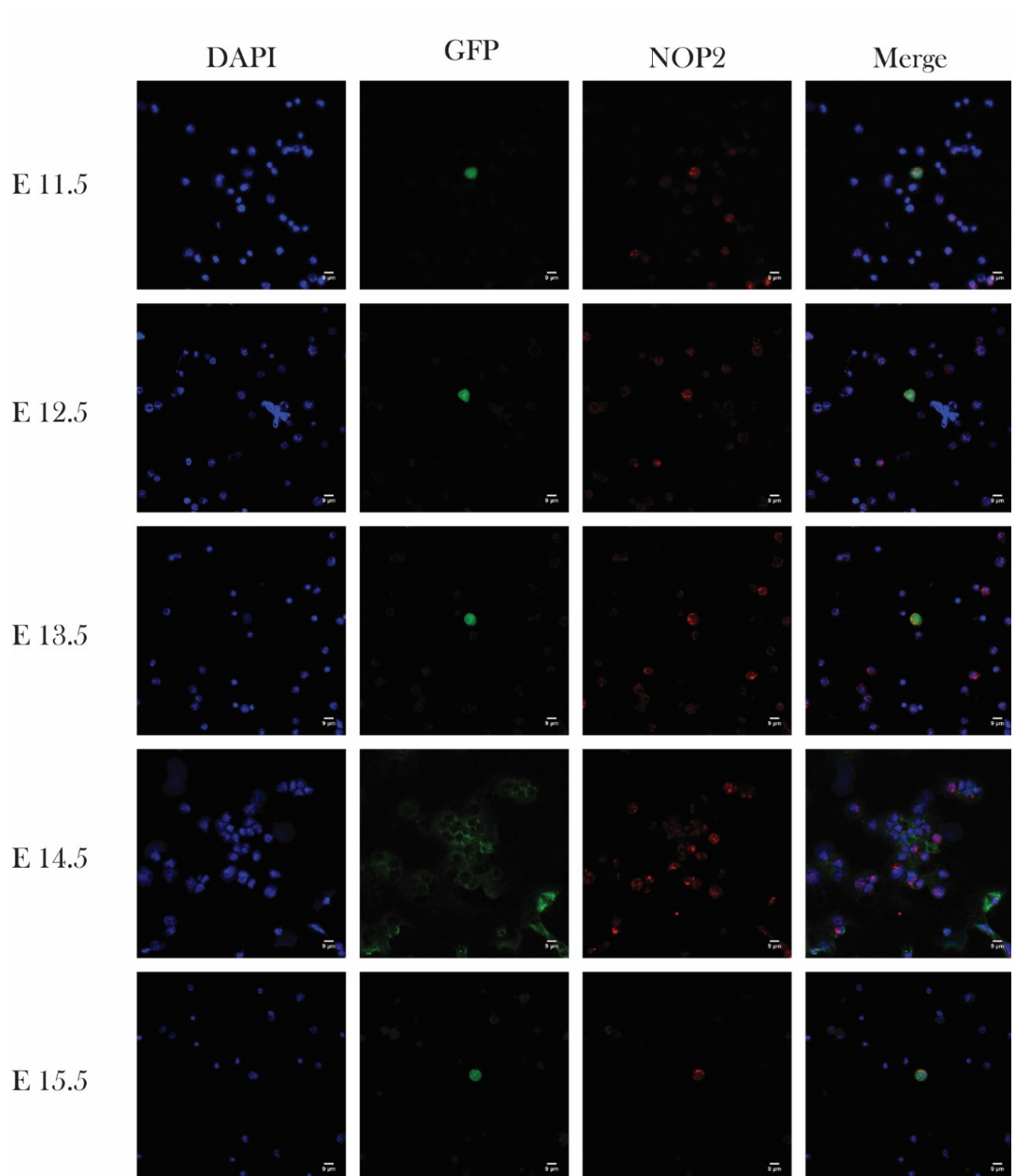


**Fig 5.3 NOP2 expression in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9 μm.

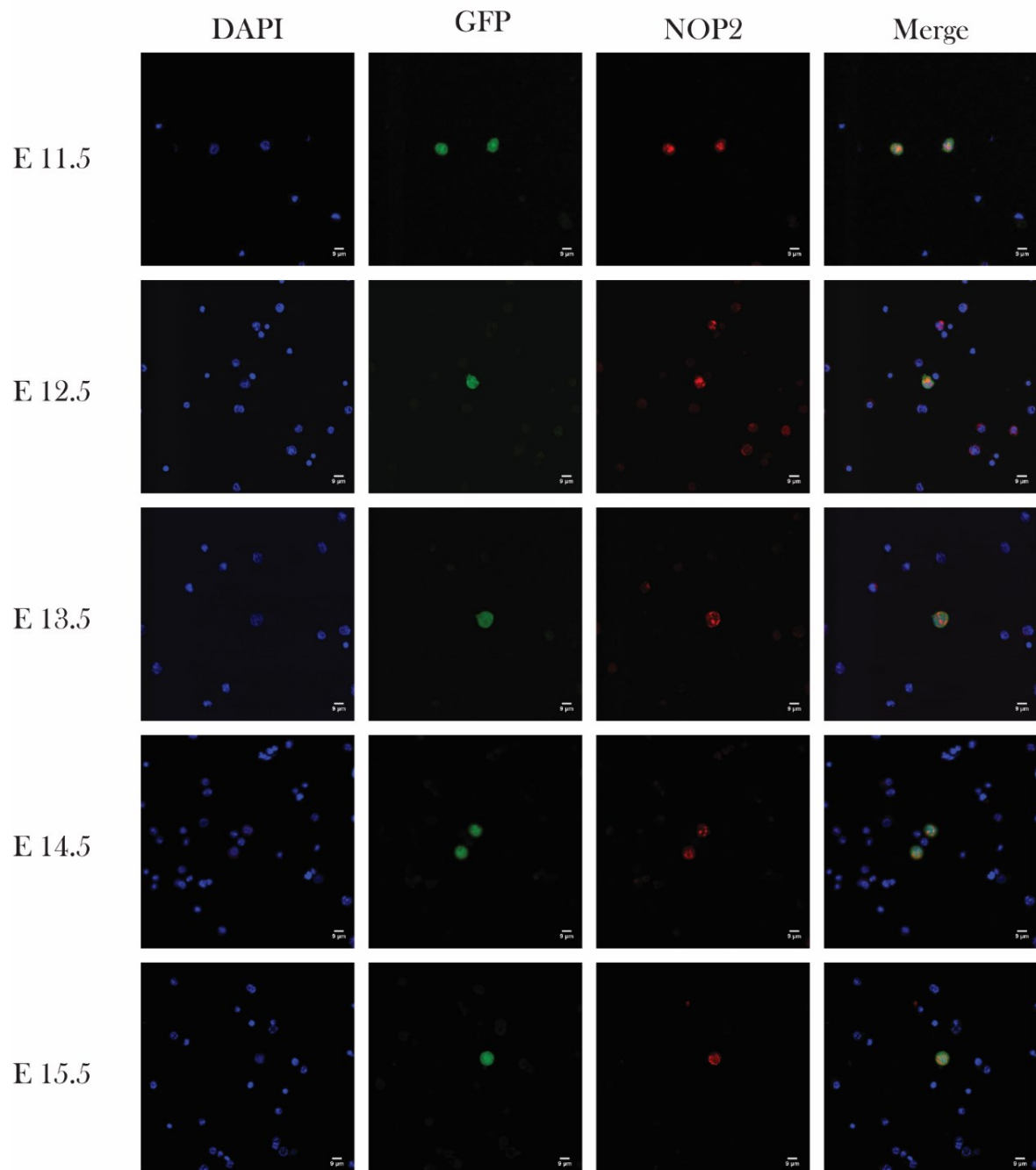


**Fig 5.4 NOP2 expression in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=3 (E11.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

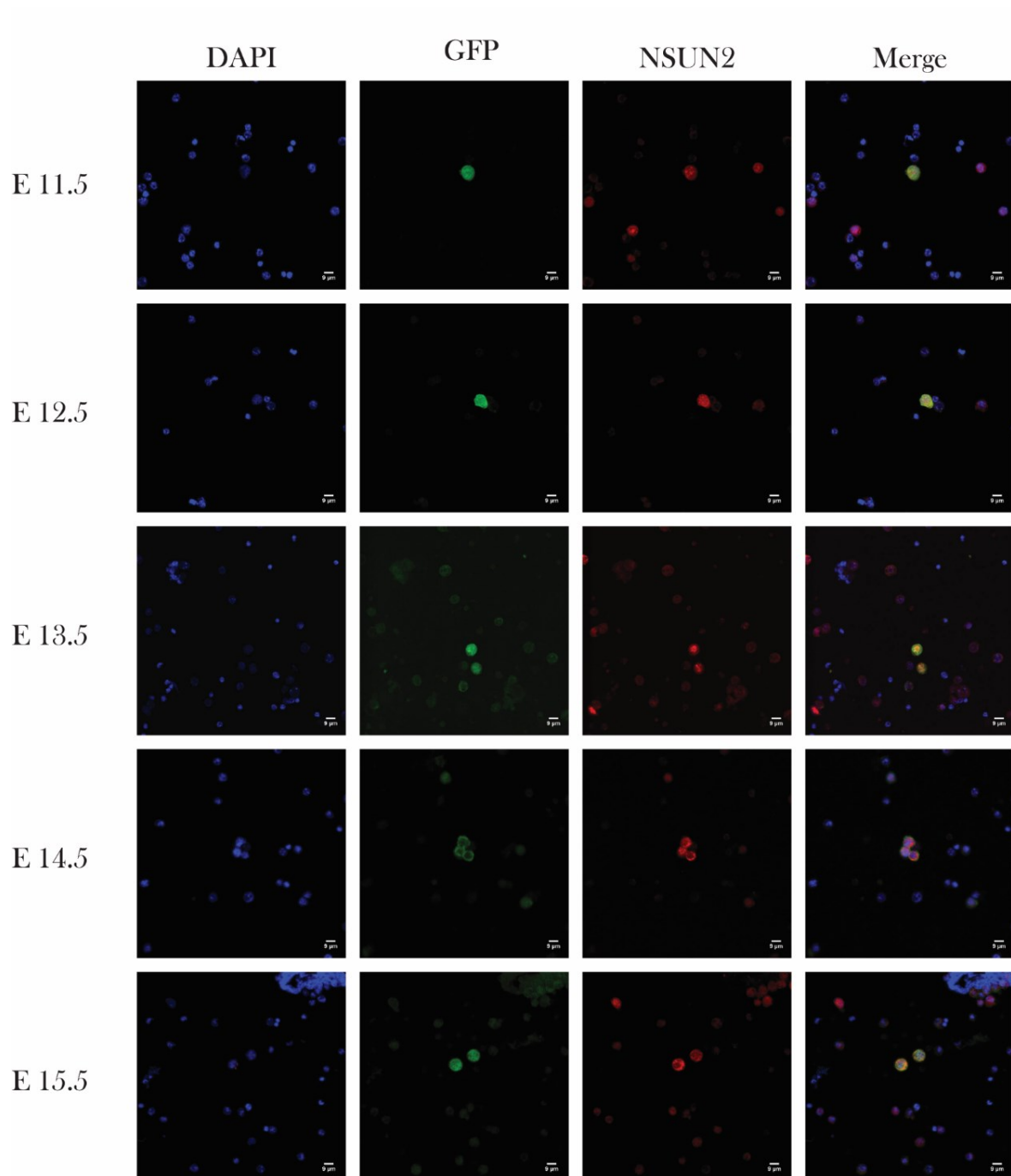


**Fig 5.5 NSUN2 expression in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E14.5) or n=3 (E11.5-E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

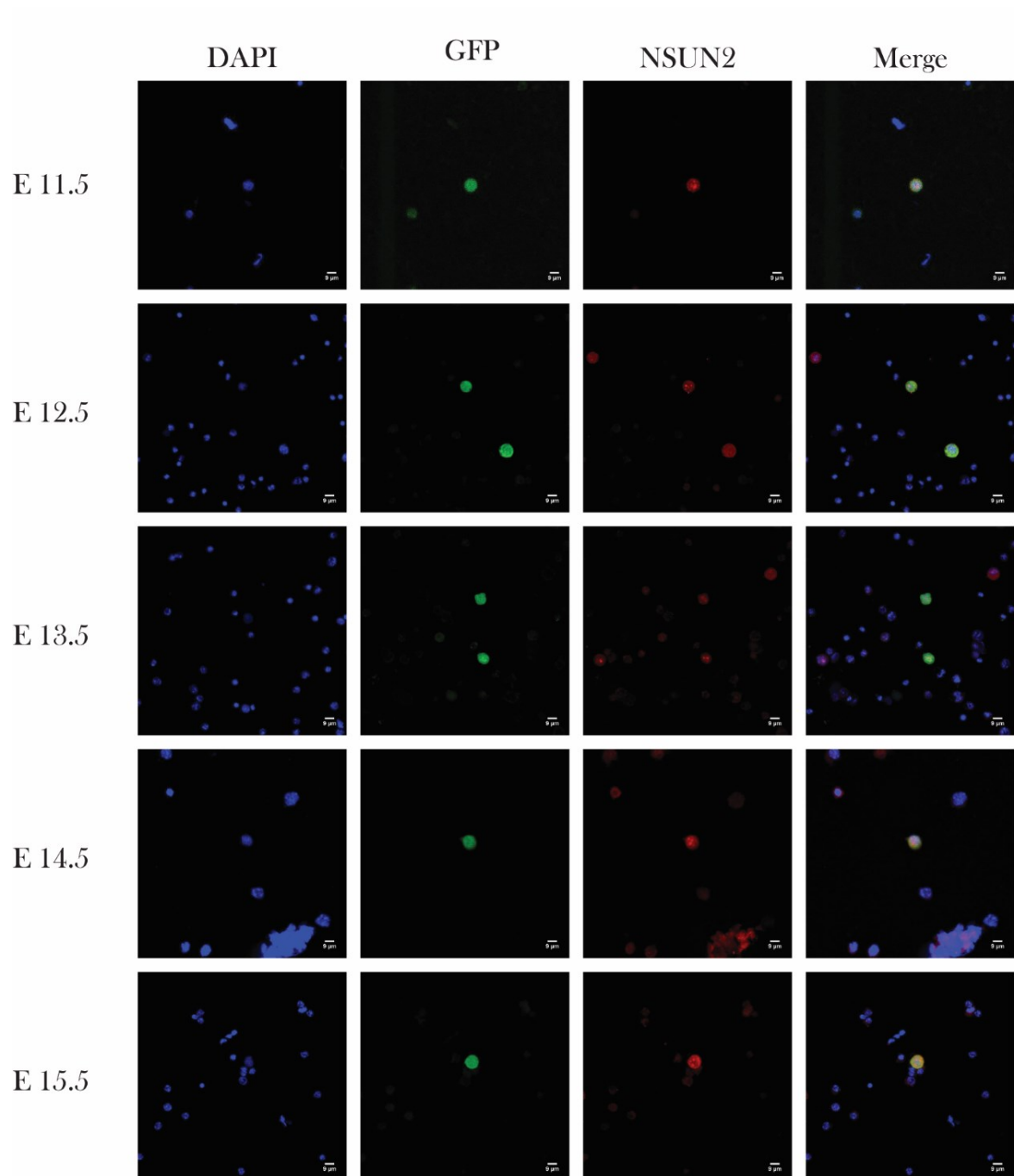


**Fig 5.6 NSUN2 expression in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9 µm.

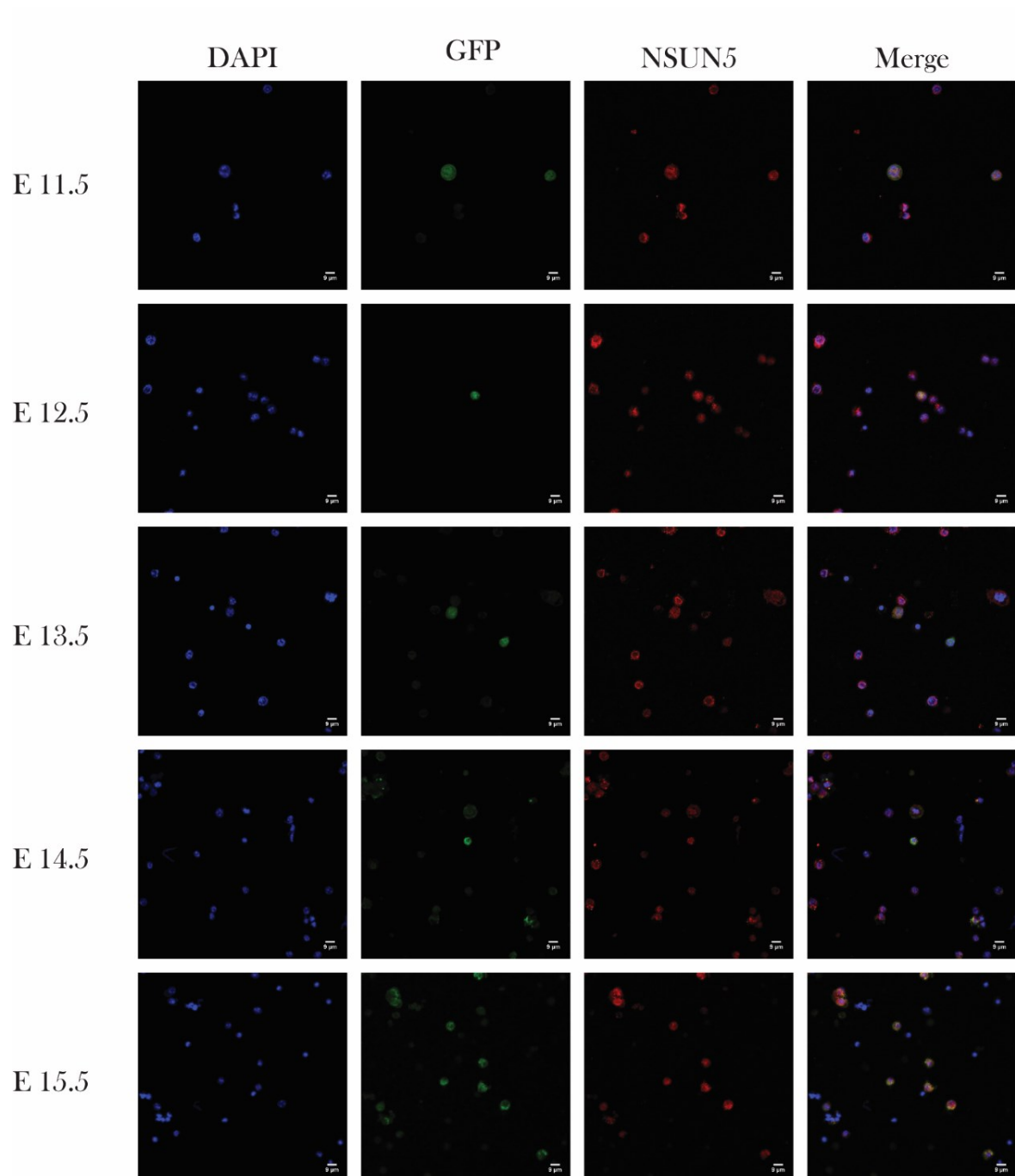


**Fig 5.7 NSUN5 expression in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5), n=2 (E14.5), or n=3 (E12.5, E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

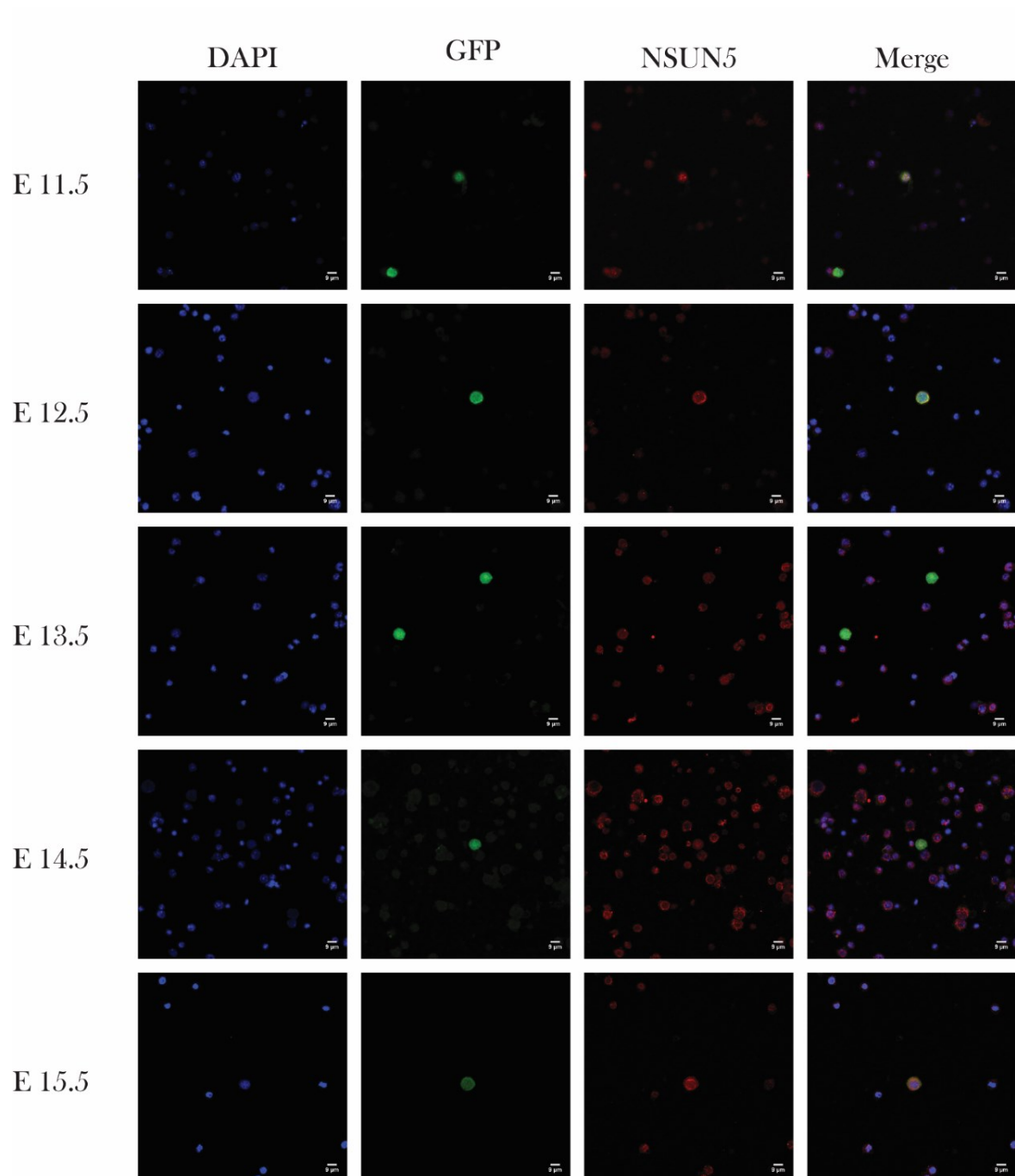


**Fig 5.8 NSUN5 expression in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu\text{m}$ .

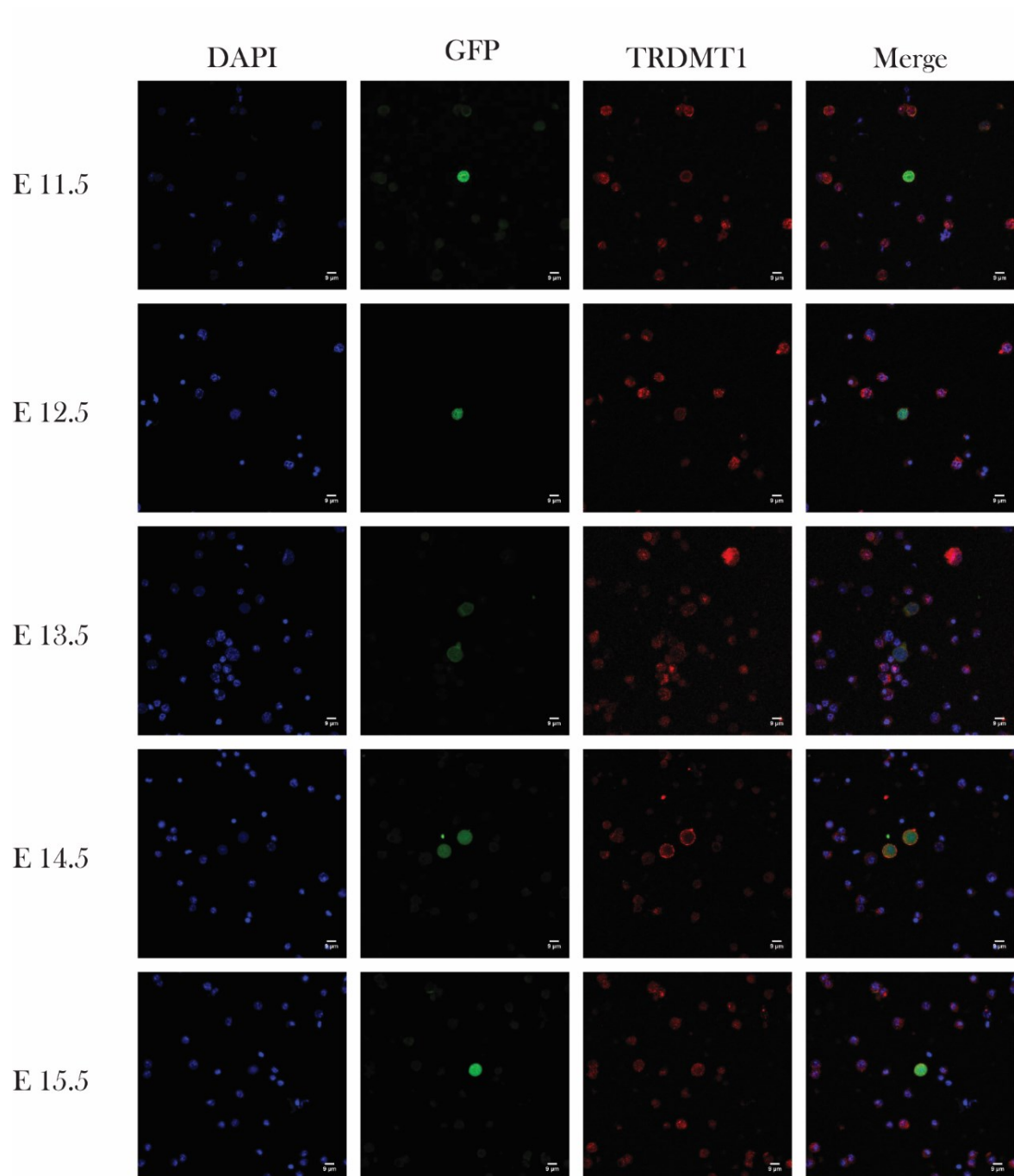


**Fig 5.9 TRDMT1 expression in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and TRDMT1. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E14.5) or n=3 (E11.5-E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9 µm.

The time points shown in figures 5.2-5.10 stretch from the rapid demethylation that occurs around E11.5 and E12.5 through to E15.5 (Hajkova et al., 2002, Hill et al., 2018). Sexually dimorphic cell programmes occur in which male PGCs undergo mitotic arrest at E13.5, while at this point the female PGCs enter into the meiotic program before themselves undergoing mitotic arrest at E14.5 (Saitou et al., 2011). These timescales are not absolute within the *Mus musculus* species, with temporal shifts observed depending upon the genetic background (P Hajkova, personal communication). The C57bl6 strain used undergo demethylation to ~E12.5-E13.5, representing a shift in time course outlined above by approximately 1 day (P Hajkova, personal communication). Consistent with this, gestation in this strain is at the later end of the range shown in mice (19-21 days). Even with this factored in, the developmental time course from E11.5-E15.5 goes beyond epigenetic reprogramming. The later time points were included due to initial observations from test stainings performed on genital ridges from embryos of these ages. These were used in this instance to provide more material for assessing staining parameters.

It can be clearly seen from figures 5.2-5.5 that there is an increased expression of NOP2, and NSUN2, in male and female PGCs relative to soma in all stages shown. However, the expression of NSUN5 in female PGCs is similar to that of the surrounding somatic cells. Intriguingly, in male PGCs NSUN5 is more highly expressed than the surrounding cells at E11.5 and E12.5, before achieving similar expression levels to somatic cells at E13.5. This data cannot be made directly comparable to the expression data from RNA-seq, however it is exciting to have a pronounced signal increase in this RNMT during *in vivo* epigenetic reprogramming. Due to time constraints and the stochastic nature of male and female frequencies within litters, TRDMT1 was only analysed in male PGCs. It showed the inverse dynamics to NSUN5, insofar as TRDMT1 displaying a lower signal in PGCs compared to soma during epigenetic reprogramming before increasing to somatic levels, or greater. These findings were encouraging, particularly as the comparison to soma was not available within the experimentation in stem cells

outlined in Chapter 4: RNA methylation in stem cells. The equivalent data is shown in figures 5.10-5.16 to show the subcellular localisation of these proteins.

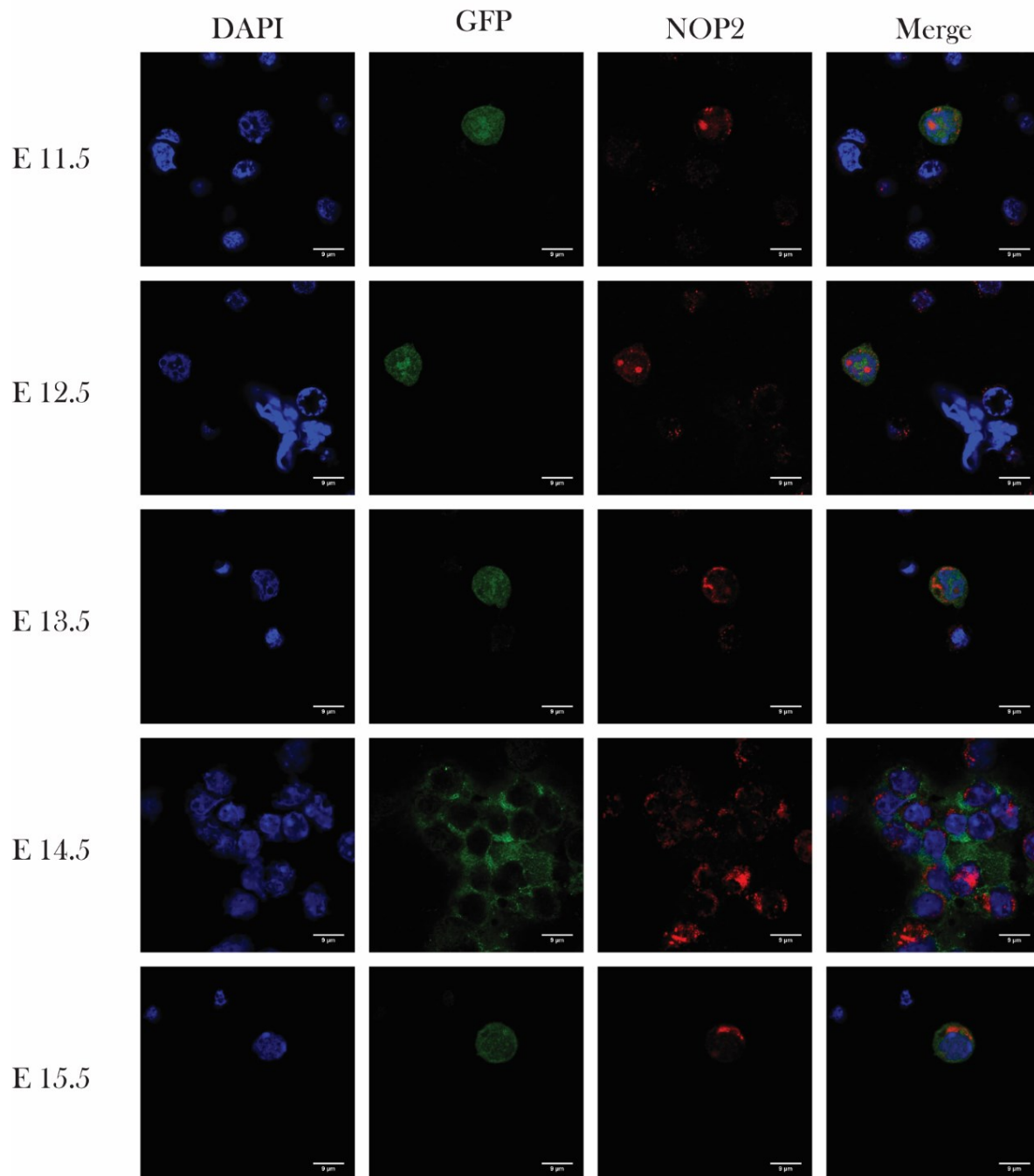


**Fig 5.10 NOP2 subcellular localisation in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9 µm.

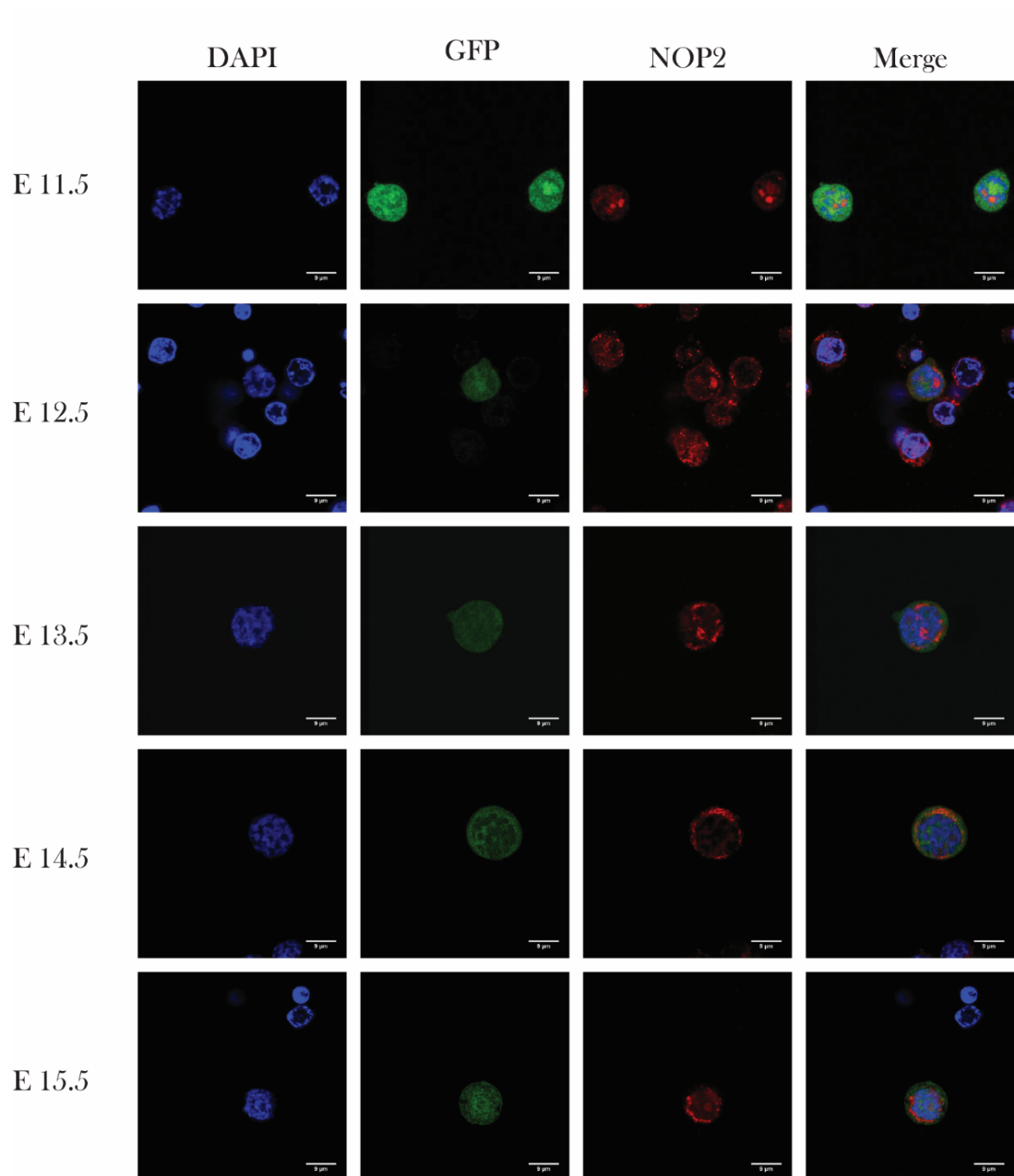


**Fig 5.11 NOP2 subcellular localisation in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=3 (E11.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu\text{m}$ .

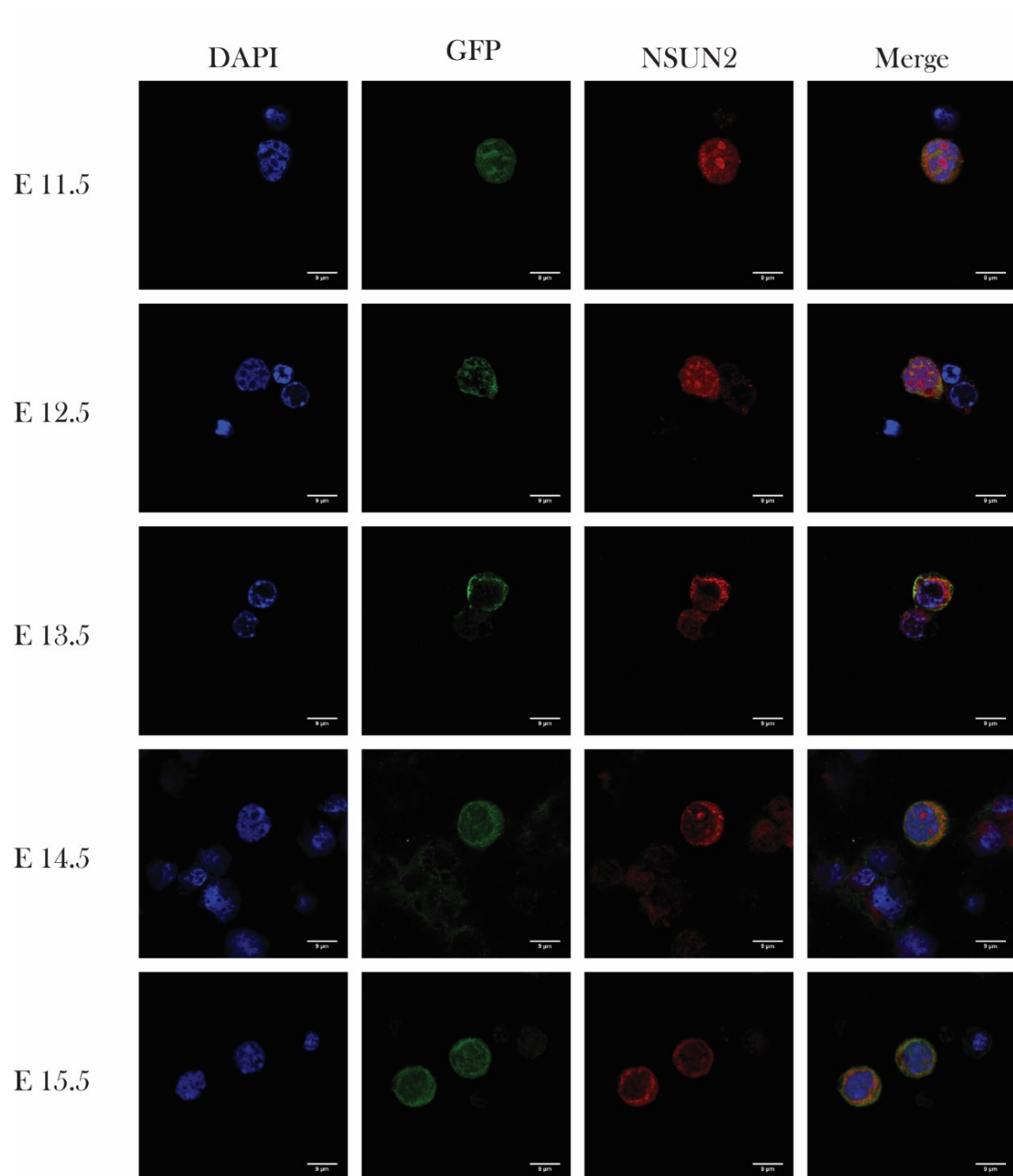


**Fig 5.12 NSUN2 subcellular localisation in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E14.5) or n=3 (E11.5-E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

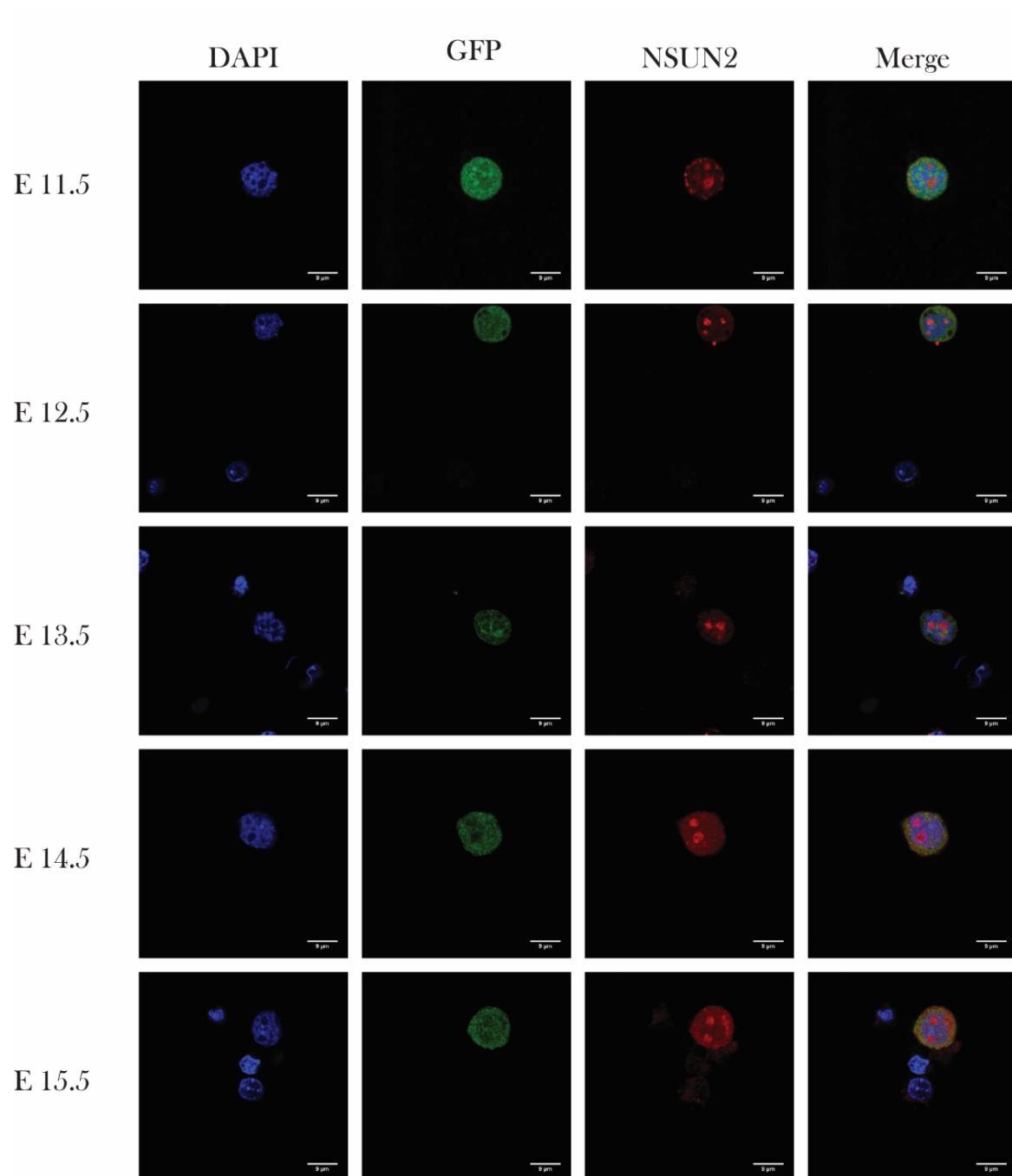


**Fig 5.13 NSUN2 subcellular localisation in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

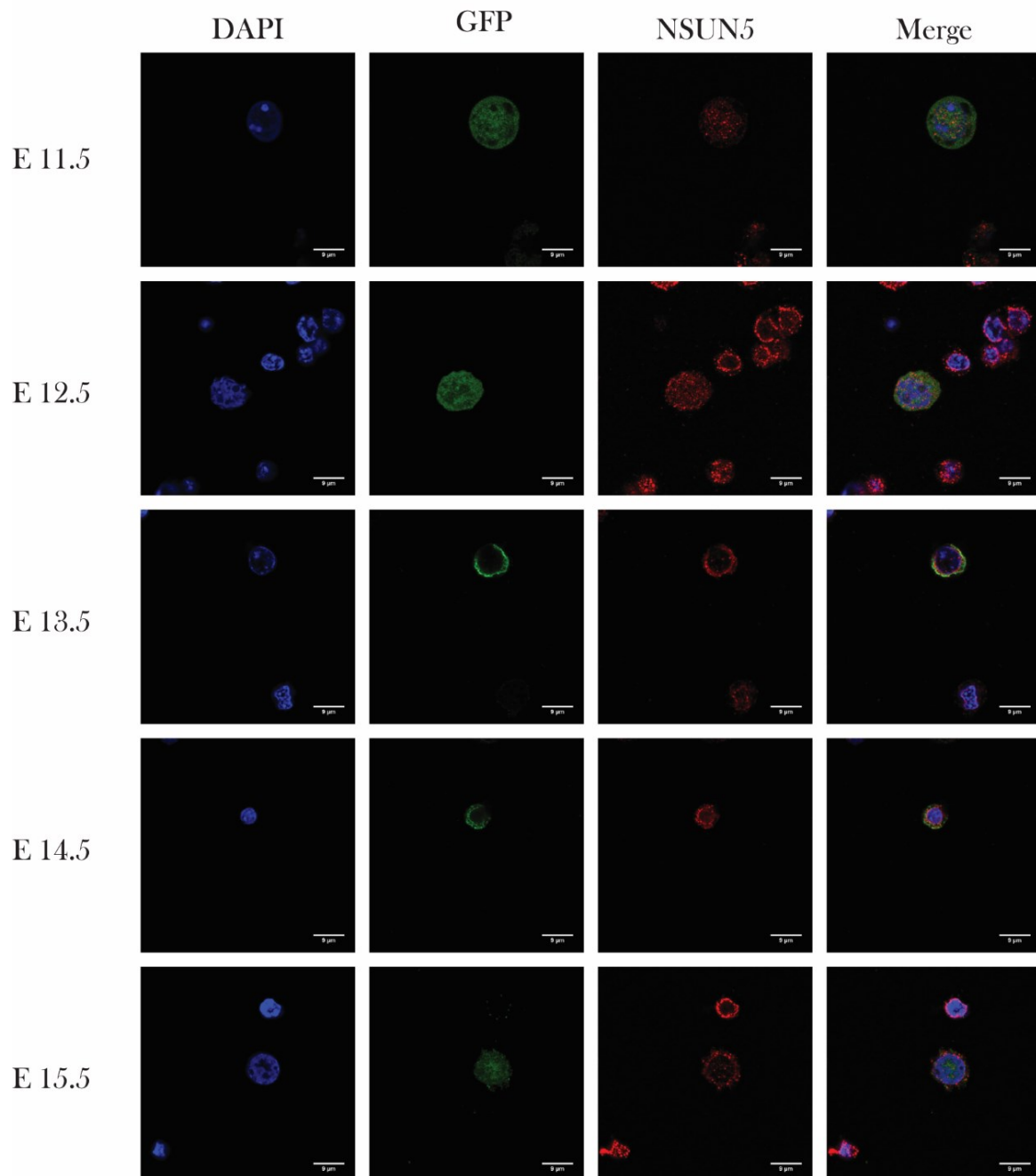


**Fig 5.14 NSUN5 subcellular localisation in female PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=1 (E11.5), n=2 (E14.5), or n=3 (E12.5, E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

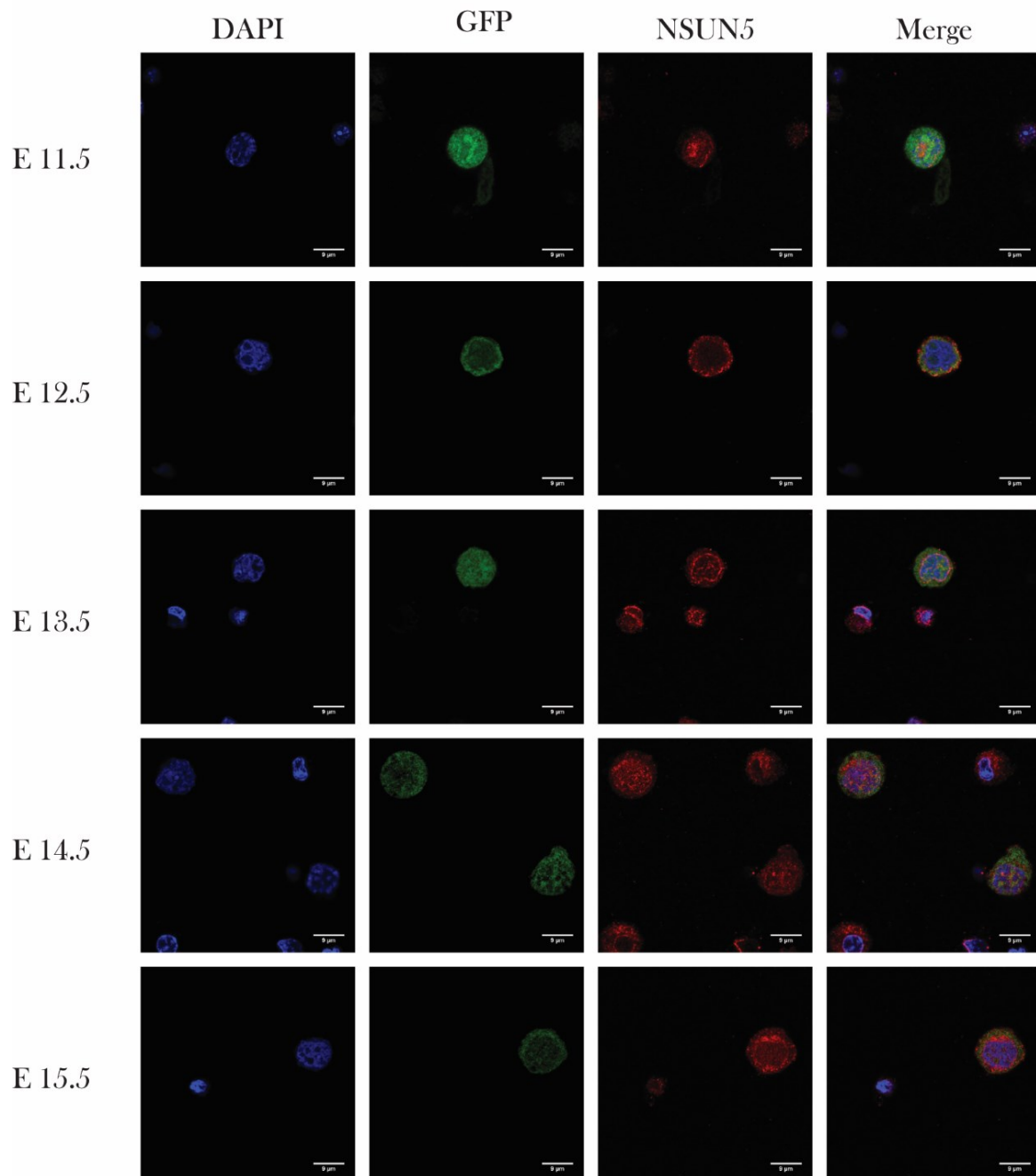


**Fig 5.15 NSUN5 subcellular localisation in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and NSUN5. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E11.5) or n=3 (E12.5-E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

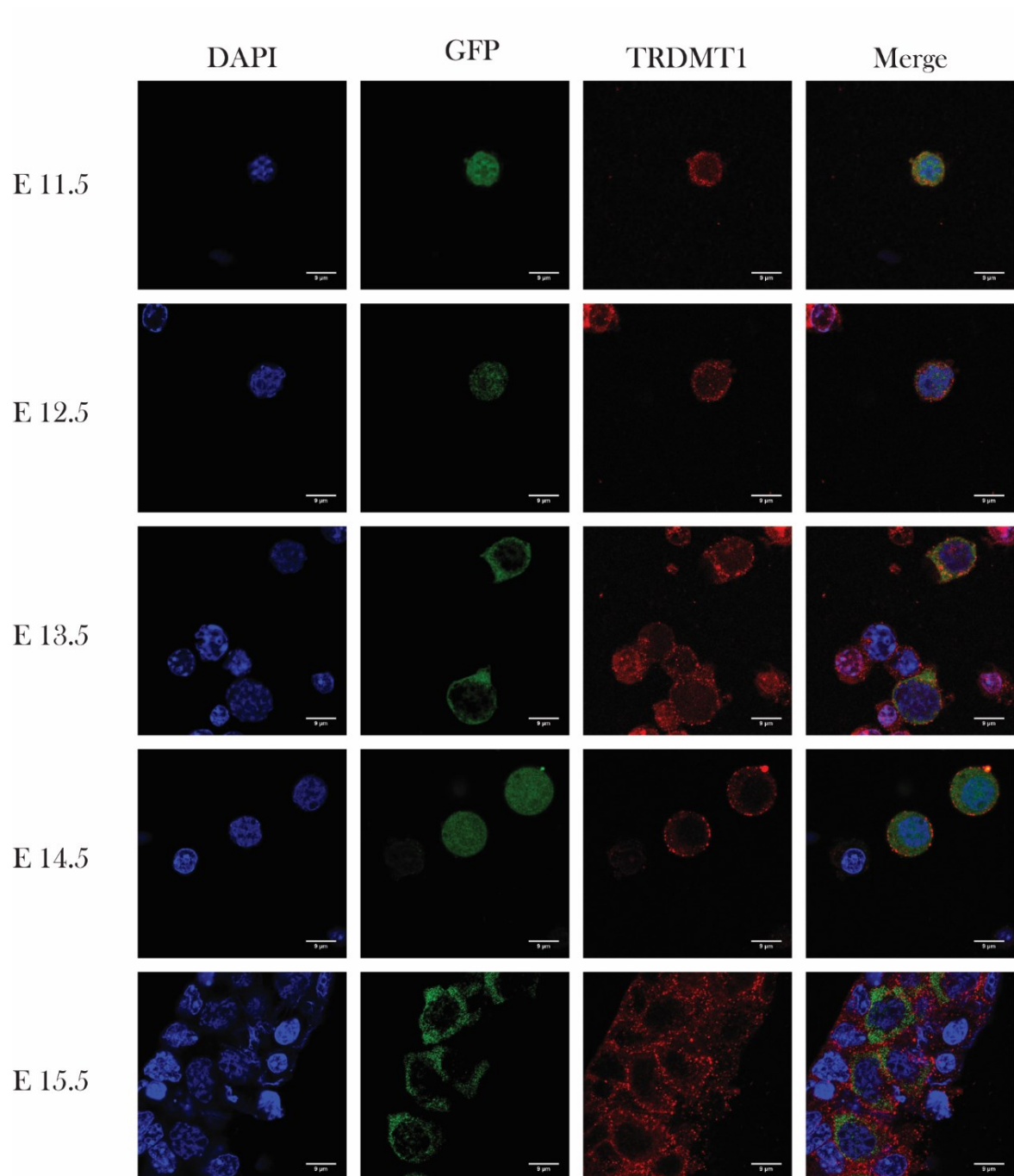


**Fig 5.16 TRDMT1 subcellular localisation in male PGCs.**

Dissociated genital ridge cells were imaged using DAPI and fluorophore-coupled antibody complexes against GFP and TRDMT1. The individual channels used to image these fluorophores are provided, as well as a merged image. C57bl6 PEOCT4GFP heterozygous embryos were used so that PGCs could be identified from soma by the expression of GFP. Representative images are displayed of n=2 (E14.5) or n=3 (E11.5-E13.5, E15.5) biological replicates with at least 5 PGCs assessed as technical replicates. Scale bars depict 9  $\mu$ m.

NOP2 initially localises to the nucleus of PGCs, with distinct nucleolar foci represented by the gaps in DAPI signalling. However, at E13.5 it redistributes to the cytoplasm, displaying a particularly strong signal around the external nuclear periphery in both males and females (Figure 5.10-5.11). Similar phenomena are observed for NSUN2 in female PGCs, although the transition occurs later at E14.5 (Figure 5.12). No such transition is observed in male PGCs, which retain a nucleolar NSUN2 signal from E11.5-E15.5 (Figure 5.13). NSUN5 shifts from a nuclear staining towards a diffuse, broadly cytoplasmic staining at E12.5 (Figure 5.14-5.15). TRDMT1 remains cytoplasmic throughout this time course (Figure 5.16). Accordingly, three distinct temporally separate RNMT subcellular localisation changes have been identified, one of which is sexually dimorphic. NSUN5 and NOP2 localise to the cytoplasm as the PGCs undergo the final stages of epigenetic reprogramming and the concurrent reduction in DNA methylation. Within the genetic background used across these experiments, E14.5 female PGCs are entering meiosis, while their male counterparts are entering mitotic arrest. It is perhaps unsurprising then that the redistribution of NSUN2 at E14.5 only occurs in one of the sexes. It was thus hypothesised that this represented a functionality related to the meiotic programme as opposed to any facet of epigenetic reprogramming. To investigate further how these RNMTs localise within cells undergoing epigenetic reprogramming, *in vivo* early embryo immunofluorescent staining was performed.

### 5.3.2.2 Early embryos

B6CBAF1/J mice were super ovulated prior to the dissection of cumulus-oocyte complexes and subsequent *in vitro* fertilisation. Dr B. Nashun completed these steps, with the downstream immunofluorescent staining, imaging, and data analysis performed by myself. There is little localisation change observed for NOP2 during this time course (Figure 5.17). In each instance a diffuse signal is observed which is excluded from the nucleoli. Up to and including the 4-cell

stage, NSUN2 primarily localises to the cell membrane (Figure 5.18). From the 4-cell stage through to the morula and blastocyst, NSUN2 signal is observed throughout the cell structure, with the exception of the nucleoli. That the strong signal at the cell membrane is not present in all embryonic stages argues against an interpretation that this may be an edging effect due to incomplete antibody permeability into the cell. Rather, this is more consistent with sequestering of the enzyme away from its substrate during the first 2 cell divisions following fertilisation. It is revelatory that the localisation changes observed for NOP2 and NSUN2 in PGCs do not coherently transpose onto the early embryo. This strongly suggests divergent molecular mechanisms, which is not surprising given that PGCs undergo active demethylation of DNA originating from both parents, while the early embryo displays passive demethylation of maternally-inherited chromosomes (Hajkova et al., 2002, Hill et al., 2018, Kafri et al., 1992, Monk et al., 1987, Howlett et al., 1991). Figure 5.19 displays the NOP2 and NSUN2 blastocyst staining at a lower magnification to enable comparison between the inner cell mass (ICM) and the surrounding trophectoderm. While there is no difference in NOP2 signalling (Fig 5.19A), NSUN2 gives a greater signal in the trophectoderm than in the ICM (Fig 5.19B). This opens up the novel context in which NSUN2 may have a specific functional role in the extraembryonic trophectodermal lineage. Further investigation would be necessary to elucidate this further. It would have been interesting to have performed similar stainings for NSUN5 and TRDMT1. However, this was unfortunately not possible due to time constraints.

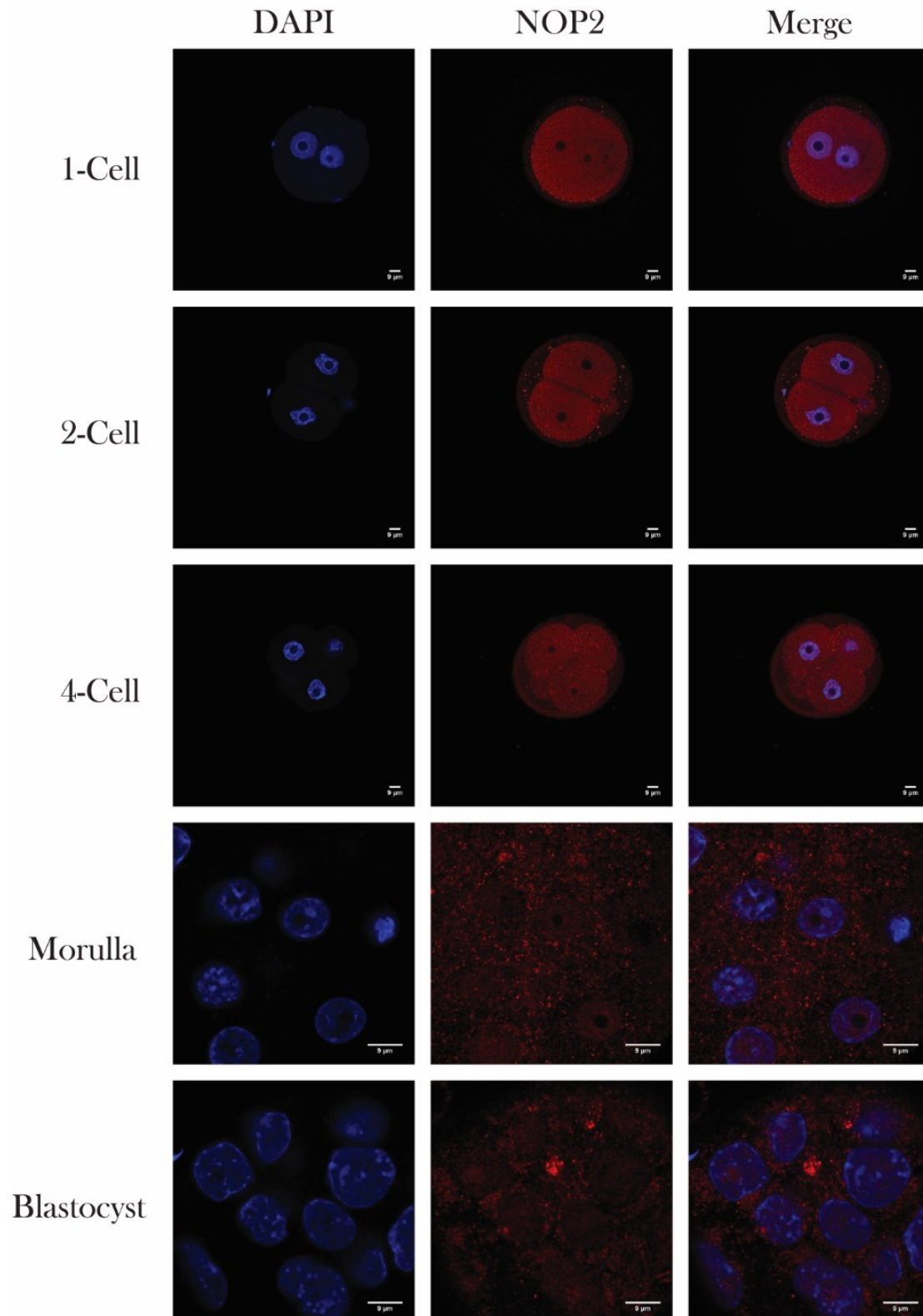


**Fig 5.17 NOP2 subcellular localisation in early embryos.**

Early embryos were imaged using DAPI and fluorophore-coupled antibody complexes against NOP2. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of n=5 (morula), n=6 (2-cell, 4-cell, blastocyst), or n=8 (1-cell) biological replicates. Scale bars depict 9  $\mu$ m.

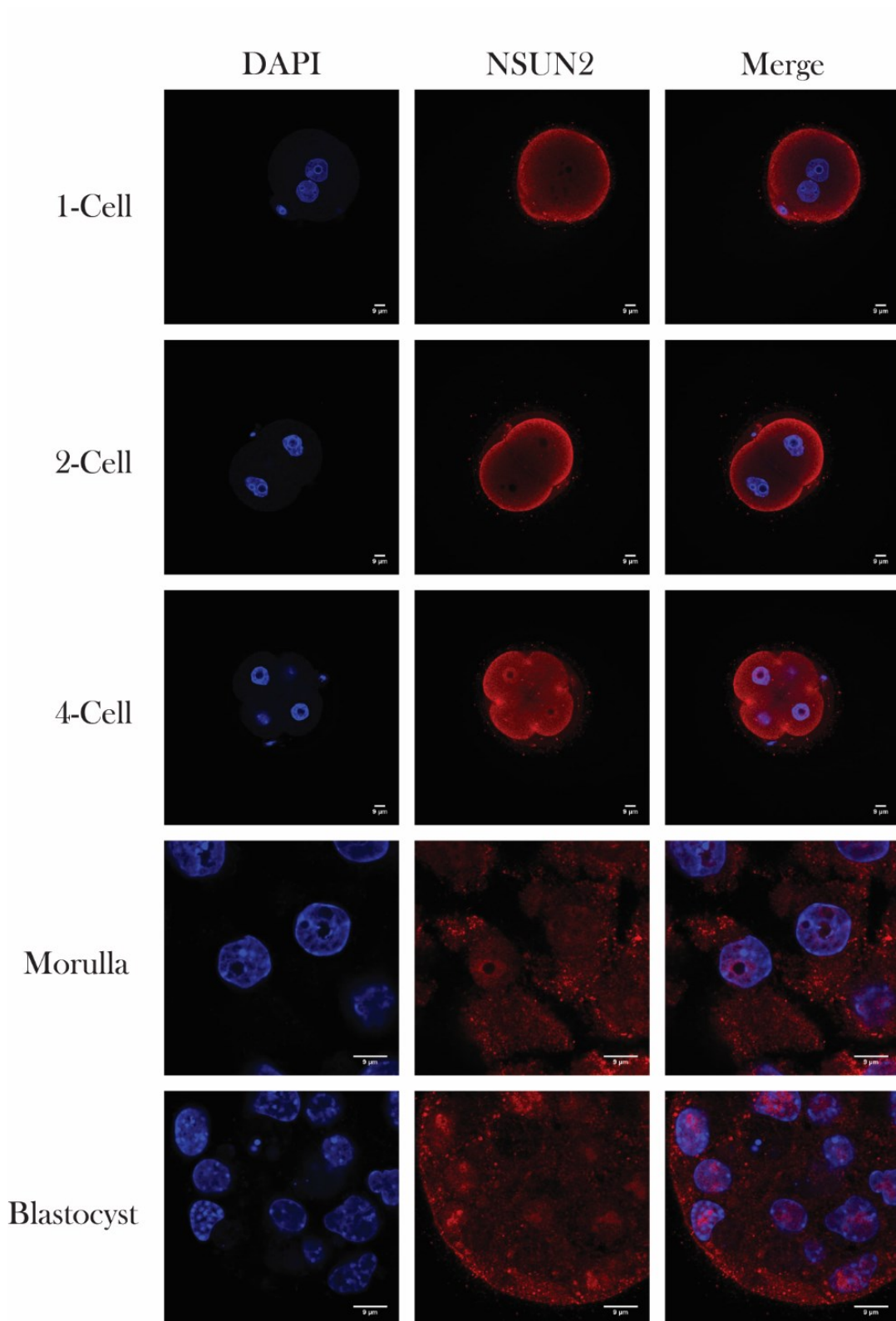


**Fig 5.18 NSUN2 subcellular localisation in early embryos.**

Early embryos were imaged using DAPI and fluorophore-coupled antibody complexes against NSUN2. The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of n=5 (morula, blastocyst), n=6 (4-cell), or n=7 (1-cell, 2-cell) biological replicates. Scale bars depict 9  $\mu\text{m}$ .

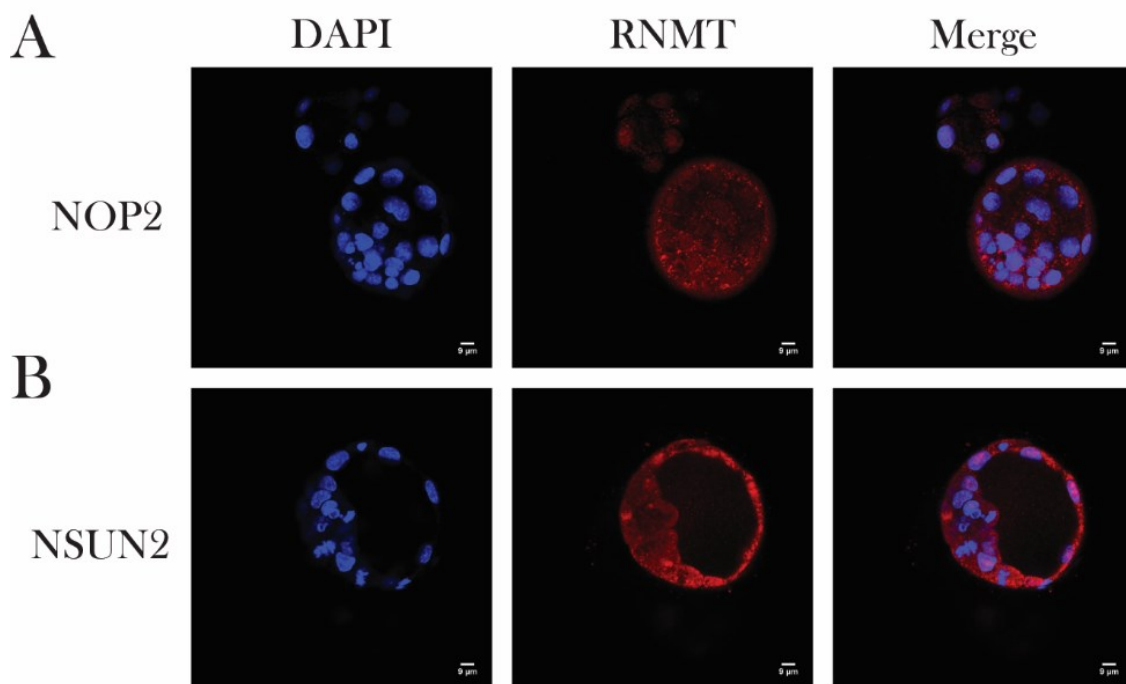


**Fig 5.19 NOP2 and NSUN2 expression in blastocysts**

Early embryos were imaged using DAPI and fluorophore-coupled antibody complexes against NOP2 (**A**) or NSUN2 (**B**). The individual channels used to image these fluorophores are provided, as well as a merged image. Representative images are displayed of n=8 (NOP2), or n=5 (NSUN2) biological replicates. Scale bars depict 9 μm.

### 5.3.3 PGC total RNA modification abundances provide tentative indications of 5mC and m6A elevations predating epigenetic reprogramming

LC-MS/MS was used to analyse the abundance of RNA 5mC, 5hmC, 5fC and m6A in primordial germ cells. C57bl6 PEOCT4GFP heterozygous embryos were used, as was the case for the immunofluorescent stainings. A pure population of PGCs was generated by FACS sorting cells dissociated from dissected embryonic genital ridges for GFP expressed under *Oct4* promoter. The GFP-negative somatic cells were collected as well as the GFP-positive PGCs. This enabled comparisons between the PGCs which were undergoing, or had underwent, epigenetic reprogramming with surrounding somatic cells. As such, any behaviour specific to the PGC biological context could be ascertained by controlling for the condition in the soma. This is in addition to assaying dynamics along a temporal dimension from E10.5-E15.5 embryos. Male

and female embryos were treated as separate samples, with the results obtained outlined in figure 5.20.

Due to the low cell numbers involved (as low as 67 PGCs in some dissections), the material available for analysis was very low. As a result of this, RNA 5hmC and 5fC could not be detected above background noise from these samples by LC-MS/MS. Independent verification by dot blotting was also not possible due to the comparatively low amount of RNA which could be recovered. RNA 5mC shows no change in any of the samples measured in either male or female embryos (Figure 5.20A-B). In female PGCs there is an upregulation in m6A from E10.5 to E12.5 when viewed with a 2-day time resolution (Figure 5.20C). This shows significance to  $p < 0.05$ , and should be considered in the context of the large variability of m6A signal from E10.5 and E11.5 PGCs. In males a similar increase is observed, but rather this is at 1-day time resolution from E10.5 to E11.5 (Figure 5.20C). When considered together, the landscape for RNA 5mC and m6A is relatively static at the total RNA level over this period of development. At the earliest bounds of the time course assayed, m6A is less prevalent in PGCs than at subsequent time points. This may point towards an earlier event of which only the tail end is caught by this experiment. The sample variability in embryos at the beginning of this time course is intriguing. E10.5 is the only sample in which embryos are pooled (due to low cell numbers), which should reduce the variability between samples. Rather, as the variation is between litters, this is more indicative of differences related to slightly different embryonic ages at the point of collection relative to fertilisation. 6-Methyladenosine thus may increase rapidly in PGCs leading up to E11.5. Applying this assay to PGCs and somatic cells from younger embryos would enable this to be elucidated further. The percentage abundance values obtained are approximately in line with those obtained from total RNA from ESCs.

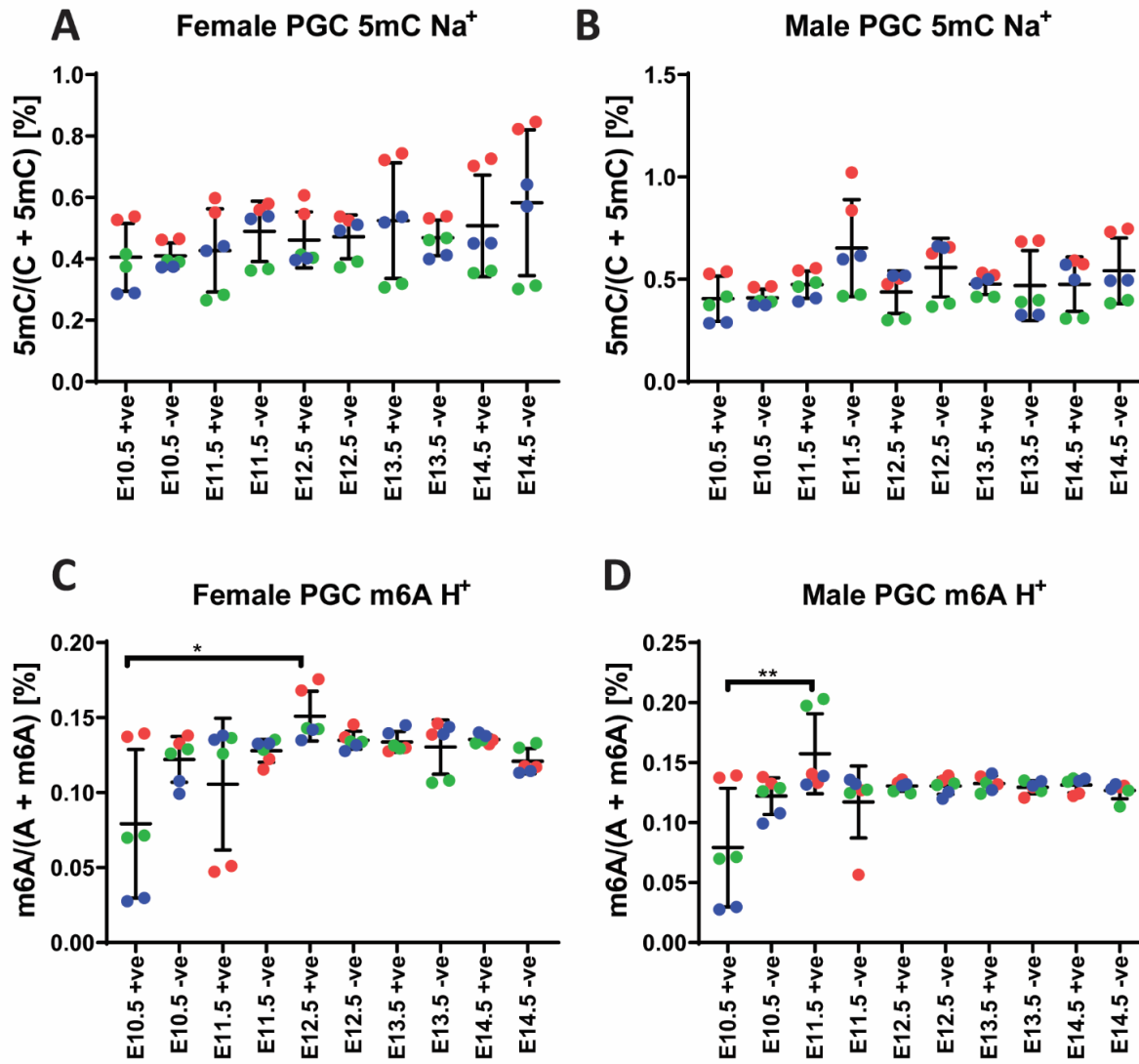


**Fig 5.20 Modified RNA quantification by LC-MS/MS in PGCs and surrounding somatic cells.**

5-methylcytosine (**A-B**) and 6-methyladenosine (**C-D**) levels in total RNA from GFP positive PGCs and GFP negative somatic cells are provided as a percentage of all quantified forms of the respective unmodified nucleoside. Female (**A, C**) and male (**B, D**) embryos are displayed separately. The age of the embryos, as well as whether the cell population were GFP-positive PGCs (+ve) or GFP-negative somatic cells (-ve), are provided on the x-axis. Technical duplicates of biological triplicates (n=3) are shown, with the technical replicates matched by colour. E10.5 samples were pooled from littermates, with the same data points plotted for male and females. All other samples represent individual embryos. Statistics were performed using two-way ANOVA with Tukey T-test for the biologically relevant comparisons in which only developmental age, or GFP expression varied. Where statistically significant differences occur, they are demarcated as follows: \* = p < 0.05, \*\* = p < 0.005, \*\*\* = p < 0.0005, \*\*\*\* = p < 0.0001.

## 5.4 Discussion

Through the development and refinement of methodologies, and their application in *in vitro* stem cell models, the work outlined in previous chapters built up to the material in this chapter. Through analysing previously published RNA-seq data, encouraging dynamics in multiple RNMTs was observed during epigenetic reprogramming in PGCs (Hill et al, 2018). However, there were no corresponding changes in 5mC levels in total RNA during, or after, reprogramming in either female or male embryos. In spite of this, three separate reproducible subcellular localisation changes were revealed for the RNMTs NSUN5, NOP2, and NSUN2. These progressively occurred at E12.5, E13.5, and E14.5 respectively, with NSUN2 only displaying this phenomenon in female PGCs. When considered together, there are strong implications that there is an underlying biological phenomena, potentially synergistically brought about by sequential relocations of RNMTs from the nucleus to the cytoplasm. Bulk averaging may be concealing biologically relevant alterations in RNA 5mC abundance within RNA subsets, or even specific transcripts. It would be interesting to perform RNA subtype selection on PGCs derived RNA, however substantial pooling of samples would be necessary. Pooling of samples was necessary for E10.5 embryos to enable quantification of 5mC and m6A abundances.

The early embryo facilitated the analysis of RNMT expression and localisation in an epigenetic reprogramming context brought about in part by a passive, as opposed to entirely active, mechanism. Within this situation *Nsun2* was found to steadily increase in expression up to the 4-cell stage, at which point the protein is relocalised from the proximity to the cell membrane, to the rest of the cell, excluding the nucleolus (B. Nashun). This is accompanied by a reduction in the expression at the mRNA level. *Nop2* undergoes a ~100-fold increase in mRNA expression from the 1-cell to 2-cell stage before precipitously decreasing as the embryo develops. No change in the subcellular localisation of this protein was observed. RNA modification quantification in

the early embryo model would be key to more fully characterising this system, particularly as the data presented points towards a dynamic, regulated system.

The readers, writers, and erasers of 6-methyladenosine are minimally altered through PGC epigenetic reprogramming. This corresponds with a relatively static m6A abundance at the total RNA level. There are more interesting expression dynamics in the early embryo. For example, the transient upregulation of YTHDF2, which is necessary to clear maternal transcripts in the maternal to zygotic transition to enable development beyond the 2-cell stage (Ivanova et al., 2017). 6-Methyladenosine quantification in early embryos would accordingly be a potentially revelatory line of inquiry.

# Chapter 6: Discussions and conclusions

Within the Introduction and Literature Review chapter of this thesis the current understandings of RNA 5mC, 5hmC, and m6A were outlined. The underlying connections linking these modifications to development, and postulating their role in epigenetic reprogramming in particular, were established. This built towards the key hypotheses outlined in that chapter and provided below:

- In the establishment of DNA 5mdC as a mechanism to regulate genome utilisation, RNA 5mC has become connected to genomic control and developmental processes. This is facilitated through an incomplete separation of the molecular machineries associated to these two evolutionarily-linked nucleoside modifications
- RNA 5mC and 5hmC have biologically relevant functions, some of which are more relevant in, or are specific to, developmental processes such as epigenetic reprogramming
- RNA 5mC and 5hmC levels change globally and locally during the time-course of epigenetic reprogramming in response to the global erasure of DNA methylation as an epigenetic regulatory mechanism
- RNA 5mC exhibits a functional interplay with 5mdC
- 6-methyladenine is involved in the cell fate transitions, and so is expected to alter in its abundance during epigenetic reprogramming

## 6.1 Technical developments

In order to address the above hypotheses and better understand the underlying biological significance and roles of these modifications, dot blotting and liquid chromatography tandem mass spectrometry techniques were developed and optimised. RNA dot blotting was initially considered as a relatively straight-forward technique, however a considerable amount of time and effort was required to advance it to the point of being applicable. This was principally due to artefacts arising from DNA 5mdC signal which were not properly controlled in the published approaches available (Mishima et al., 2015, Hodge, 1998, Miao et al., 2016, Cui et al., 2017). Wholesale changes were thus needed to enable valid results to be generated from this technique.

This casts doubt over the interpretation of RNA dot blotting which is not fully controlled for other nucleoside contaminants, principally in relation to the 33D3 anti-5mC antibody. DNA 5mdC is typically more abundant in DNA than 5mC is in RNA. Additionally, the comparative amounts of each nucleoside necessary to achieve a detectable signal indicated that the antibody shows a greater affinity for 5mdC than 5mC. This exasperates the effects any contaminants could have, further highlighting the importance of sufficiently controlling for contaminant-derived signal.

Through the use of reproducible analyte retention times and two separate m/z filters, LC-MS/MS provided a highly specific methodology to quantify RNA modification abundances. Critical to this specificity is the selection of the mass of the analyte before and after reproducible fragmentation. As the glycosidic bond between the ribose and base is broken in fragmentation with the charged base detected, the mass difference is that of the ribose ring for each RNA nucleoside. As such, there is no chance of DNA contamination giving an artificial signal. The LC-MS/MS approach is a truly quantitative method, while the uncharacterised antibody response to epitope abundance results in dot blotting being at most semi-quantitative. An unparalleled level of sensitivity for the detection of modified RNA nucleosides was achieved, with detection down to  $5 \times 10^{-17}$  mol and quantification typically around  $2.5 \times 10^{-16}$  mol. This new technical set-up opens up modified RNA quantification to entirely novel low cell number contexts. This was necessary to quantify RNA 5mC and m6A in PGC RNA, although it was not sufficient for the detection of 5hmC or 5fC. This reflected the very low proportional abundance of these modifications within a starting material which was already limited by the low number of cells available. With this in mind, it is clear that, although the body of work represents a significant advancement within the field, there are biological rationales for seeking even greater sensitivity. This may not be technically feasible, and certainly is not based on currently available instrumentation for some quantified modifications. Ion suppression reduces the 5hmC signal

detected, as it co-elutes with unmodified cytidine. There is this scope to further develop the chromatography program to fully resolve these analytes to increase the sensitivity for 5hmC.

The LC-MS/MS is compatible with application to the myriad of other RNA modifications which have been discovered *in vivo*, but remain poorly characterised. Through cycling Q1 and Q3 values many analytes can be analysed in parallel without reducing the sensitivity of the assay to any of the analytes. As the sensitivity achieved in this thesis is unparalleled to date, this method opens up novel avenues of research into these modifications.

## 6.2 Quantification of modified RNAs

Upon the analysis of both *in vitro* and *in vivo* epigenetic reprogramming models, key hypotheses 1, 4, and 5 indicated on the previous page were refuted. RNA 5mC was found to be independent of DNA 5mdC level in ESC and PGC demethylation systems. This disproved their proposed functional independence and the anticipated global changes in 5mC during epigenetic reprogramming. While 5hmC could not be detected in PGCs, there was no evidence from ESCs for global, transcriptome-wide 5hmC alterations during epigenetic reprogramming. 6-Methyladenosine was muted in its response to any of the differing cell states assayed both *in vitro* and *in vivo*. No differences were observed, save for a reduction in mRNA from TKO EpiLSCs compared to wild type EpiSCs and a somewhat lower m6A levels in E10.5 PGCs than in later time points. The former is most likely a result of the TKO EpiLSC population partially entering differentiation, as was found through RT-qPCR analysis of marker expression, and so having already cleared the transcripts associated with the earlier cell fate. Further work on embryos earlier in development may identify that this is the tail end of a biologically relevant modification function. This would be consistent with the novel signalling encountered by migratory PGCs entering the gonadal niche resulting in the clearing of transcripts from the previous cell fate (Introduction and literature review: Figure 1.7) (Wang et al., 2015, Sheth et al., 2003, Du et al.,

2016, Roundtree et al., 2017). Future work will look to establish this through the collection of earlier migratory PGCs for FACS and LC-MS/MS analysis.

While the systems did not behave as had been hypothesised, the experimental approaches used did yield novel findings. It was established that embryonic stem cells lacking detectable DNA methylation are capable of developing further than had been previously believed and adopt an epiblast-derived stem cell-like fate (Sheaffer et al., 2014). In generating this new cell state, it was possible to elucidate the modified RNA landscape in the transitions from naïve pluripotency, through capacitation, and ending in a primed state in both wild type and DNMT TKO stem cells. While there is not an intrinsic link between 5mC in RNA and DNA methylation, a transient downregulation was established along the naïve-primed axis. Serum-cultured stem cells displayed lower RNA 5mC abundances than cells grown in 2i or EpiSC/EpiLSC. This was statistically significant, save for the difference between 2i and serum-cultured ESCs in wild type cells. A reciprocal pattern was observed for 5hmC, but only in small RNA. While 5hmC has been observed in RNA previously, the reciprocal relationship uncovered in this work provides the first evidence *in vivo* of 5hmC acting as a RNA demethylation end product. While RNA 5mC is an obligate precursor to 5hmC, it had been unclear whether either the amount of 5hmC, or the metabolic flux of its formation, was sufficient to have an appreciable effect on 5mC. Curiously, it is in separate data sets that the respective changes in 5mC in total RNA and 5hmC in small RNA are observed in relation to the naïve to primed developmental axis. This validates hypothesis 2 in relation to RNA methylation being linked to developmental progression. Future work will move beyond EpiLSCs to attempt to generate DNMT TKO embryoid bodies. This will be assessed *in vivo* as well by injecting TKO ESCs into wild type blastocysts to attempt to generate a chimera. To what extent TKO cells could contribute to the differing lineages would then be assessed by lineage tracing.

RNA subtype was found to be the most critical factor in establishing the abundance of 5mC, 5hmC, 5fC, and m6A, which two distinct sub-categories emerging. The modified cytosines exist at much higher frequencies in small RNAs and are relatively devoid in mRNA, while the converse is true for m6A. This is a product of the divergent deposition machineries which bring about these modifications *in vivo*. It highlights the context-dependent nature of such modifications as well as their functional and regulatory divergence. It has recently been shown The RNA m6A demethylase FTO shows different substrate affinity to nuclear and cytoplasmic mRNA (Wei et al., 2018). As such, future work will compare the abundance of RNA modifications in distinct nuclear and cytoplasmic RNA fractions. The quantified abundance of a modified RNAs does not give sequence context, and so cannot fully characterise RNA modifications *in vivo* in isolation. RNA bisulphite sequencing will be used to characterise the location of RNA 5mC in the transcriptomes of ESCs cultured in 2i, serum, and EpiSCs. Critically the transcripts which show different methylation along this naïve-primed developmental axis will be assessed to more fully understand the dynamics of RNA 5mC. This assay will be repeated with *Nop2* knockout and *Nsun2* knockout embryonic stem cells to determine the 5mC sites which are dependent upon these RNMTs.

### 6.3 TET-independent formation of RNA 5hmC and 5fC

A TET independent mechanism for the formation of 5hmC and 5fC was found to exist, through the detection of these modifications in stem cells with no TET activity. It was naturally assumed from the *in vitro* activity of TET enzymes in generating RNA 5hmC from 5mC that these enzymes were responsible for generating they higher oxidative derivatives of 5mC *in vivo* (Fu et al., 2014). That no difference in 5fC abundance was detected upon the ablation of the TET enzymes goes beyond the implication of a divergent pathway, showing the TET enzymes to be irrelevant to 5fC abundance. This was an exciting and unexpected avenue opened up by this

research and opens up fertile ground for further investigation. Alternatively, the 5hmC nucleosides detected may represent previously hydroxylated cytidines which have been retained in cellular RNA pool through nucleotide salvage (Zauri et al., 2015). While an intriguing idea, the number of passages separating the TET triple knockout formation and collection for experiments would result in much lower levels of 5hmC than are detected by LC-MS/MS. Having shown that DNA methylation is dispensable in generating an Epiblast-like stem cell fate, it would be intriguing to perform the same transformation with TET triple knockout ESCs. Combining TET triple knockout on top of the DNMT triple knockout cells would generate a novel system which truly disentangles the effects of the TET enzymes on RNA from the activity of these enzymes on DNA.

## 6.4 Subcellular localisation

Prior to this work, it had been found, or predicted, that all RNMTs are localised to the nucleolus, with the exceptions of TRDMT1 and NSUN4, which are cytoplasmic and mitochondrial respectively (Metodieiev et al, 2014, Lewinska et al., 2015). Through *in vitro* and *in vivo* PGC and early embryo systems subcellular localisation changes were observed for a number of RNMTs, revealing a system of regulation which had not been considered previously. Within the developmental progression modelled *in vitro* from naïve cells through to primed, NOP2 and NSUN2 were each found to redistribute from the cytoplasm to the nucleus, and in particular the nucleolus. For NOP2, this happens between serum-cultured cells and primed EpiSC/EpiLSCs, corresponding with an increase in RNA 5mC. As the nucleolus is major site of RNA processing, this is highly indicative of NOP2 localisation to this site mechanistically driving this increase in RNA 5mC (Hernandez-Verdun et al., 2010). While no changes in RNA 5mC were observed during, or immediately following, epigenetic reprogramming in PGCs, subcellular localisation changes for NSUN5, NOP2 and NSUN2 were observed. Within the context of the C57bl6

genetic background of the mice used, NSUN5 and NOP2 were found to relocalise from the nucleus to the cytoplasm in the later stages of DNA demethylation. NSUN2 subsequently underwent this transition only in female PGCs while undergoing meiosis, however this was not the case for male PGCs, which at this stage experience mitotic arrest. That NOP2 and NSUN2 localise to the cytoplasm in naïve cell states *in vitro* and *in vivo* is encouraging. This reinforces their regulated association with the nucleus, and so their substrate RNAs, in cells which are primed to differentiate or have differentiated. NSUN5 is cytoplasmic in each *in vitro* stem cell fate assayed, which shows that the nuclear localisation in E11.5 PGCs is not recapitulated. As such, the trigger has not been identified, but it seemingly lies further towards a differentiated state, as NSUN5 relocalisation occurs earlier than NOP2 in PGCs, which is itself recapitulated in the primed stem cell state. From the images generated signal integration will be used to enable protein quantification and comparisons of RNMT expression. The nucleolar signal intensity shall also be calculated separately to enable the nucleolar:nucleus:cytoplasm ratio of protein abundance to be determined and compared. This would enable a more quantified assessment of subcellular localisation change.

Cells experiencing passive demethylation of maternally-derived DNA by replicative dilution in the early embryo display divergent molecular mechanisms from those elucidated in PGCs undergoing active DNA demethylation. Within this context, *Nop2* dramatically increases in expression by ~100-fold between the 1-cell and 2-cell stages, while *Nsun2* displays a transient upregulation in expression which maps to a pattern of redistribution of the protein from the cell membrane at the point of inflection. Additional experimentation would be necessary to explore the function of these enzymes within this alternate demethylation system, however the available results are strongly indicative of a regulated functionality. Separately, it was found that in the blastocyst NSUN2 is enriched in the extraembryonic trophoctodermal lineage relative to the inner cell mass. From the images generated signal integration will be used to enable protein

quantification to verify this finding. This will be cross-referenced with existing RNA-seq data (Boroviak et al., 2014) This indicates that there may be a specific role for this RNMT in the development of the extraembryonic structures. This was an unanticipated finding, which opens up the possibility that NSUN2 may be a key gene in the trophectodermal lineage. To investigate this further, *Nsun2* knockout ESCs will be generated and injected into 2-cell embryos to generate chimeras. Lineage tracing will be used to determine if *Nsun2* knockout biases cell fate against the trophectoderm. The global underdevelopment of *Nsun2* knockout embryos could be mechanistically explained by a defect in extraembryonic tissue resulting in impaired embryo growth.

In summation, while a number of the central hypotheses that guided my development of the outline of this work were refuted, a range of novel and exciting findings have been elucidated while improved methodologies have opened up novel avenues for scientific inquiry.

## Bibliography

- Acampora, D., Omodei, D., Petrosino, G., Garofalo, A., Savarese, M., Nigro, V., Di Giovannantonio, L. G., Mercadante, V. and Simeone, A. (2016) 'Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast', *Cell Reports*. Cell Press, 15(12), pp. 2651–2664.
- Alexandrovich, A., Arno, M., Patient, R. K., Shah, A. M., Pizzey, J. A. and Brewer, A. C. (2006) 'Wnt2 is a direct downstream target of GATA6 during early cardiogenesis', *Mechanisms of Development*. Elsevier, 123(4), pp. 297–311.
- Anderson, R., Copeland, T. K., Schöler, H., Heasman, J. and Wylie, C. (2000) 'The onset of germ cell migration in the mouse embryo', *Mechanisms of Development*. Elsevier, 91(1–2), pp. 61–68.
- Andrews, P. D., Knatko, E., Moore, W. J. and Swedlow, J. R. (2003) 'Mitotic mechanics: the auroras come into view', *Current Opinion in Cell Biology*, 15(6), pp. 672–683.
- Amort, T., Rieder, D., Wille, A., Khokhlova-Cubberley, D., Riml, C., Trixl, L., Jia, X.-Y., Micura, R. and Lusser, A. (2017) 'Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain', *Genome Biology*, 18(1), p. 1. doi: 10.1186/s13059-016-1139-1.
- Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P. W. S., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E., Encheva, V., Kudo, N. R., Koseki, H., Sasaki, H. and Hajkova, P. (2016) 'De novo DNA methylation drives 5hmC accumulation in mouse zygotes', *Nature Cell Biology*. Nature Publishing Group, 18(2), pp. 225–233.
- Aravin, A. A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K. F., Bestor, T. and Hannon, G. J. (2008) 'A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice', *Molecular Cell*. Cell Press, 31(6), pp. 785–799.
- Arcari, P. and Hecht, S. M. (1978) 'Isoenergetic Hydride Transfer A REVERSIBLE tRNA MODIFICATION WITH CONCOMITANT ALTERATION OF BIOCHEMICAL PROPERTIES\*\*', *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 253(22), pp. 8278–8284.
- Artus, J., Piliszek, A. and Hadjantonakis, A.-K. (2011) 'The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17.', *Developmental biology*. NIH Public Access, 350(2), pp. 393–404.
- Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M. P., Cording, A. C., Doebley, A.-L., Goldstein, L. D., Lehrbach, N. J., Le Pen, J., Pintacuda, G., Sakaguchi, A., Sarkies, P., Ahmed, S. and Miska, E. A. (2012) 'piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans.', *Cell*. Elsevier, 150(1), pp. 88–99.

- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkschlager, M. and Fisher, A. G. (2006) 'Chromatin signatures of pluripotent cell lines', *Nature Cell Biology*. Nature Publishing Group, 8(5), pp. 532-538.
- Bantis, A., Giannopoulos, A., Gonidi, M., Liossi, A., Aggelonidou, E., Petrakakou, E., Athanassiades, P. and Athanassiadou, P. (2004) 'Expression of p120, Ki-67 and PCNA as proliferation biomarkers in imprint smears of prostate carcinoma and their prognostic value', *Cytopathology*, 15(1), pp. 25-31.
- Batey, J. H. (2014) 'The physics and technology of quadrupole mass spectrometers', *Vacuum*. Pergamon, 101, pp. 410-415.
- 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. and Chang, H. Y. (2014) 'm(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells.', *Cell stem cell*. Elsevier, 15(6), pp. 707-19.
- Bazett-Jones, D. P., Côté, J., Landel, C. C., Peterson, C. L. and Workman, J. L. (1999) 'The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains.', *Molecular and cellular biology*. American Society for Microbiology Journals, 19(2), pp. 1470-8.
- Belicard, T., Jareosettasin, P. and Sarkies, P. (2018) 'The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress', *BMC Biology*. BioMed Central, 16(1), p. 103.
- Beyersmann, D. (2000) 'Regulation of mammalian gene expression.', *EXS*, 89, pp. 11-28.
- Bjork, G. R., Ericson, J. U., Gustafsson, C. E. D., Hagervall, T. G., Jonsson, Y. H. and Wikstrom, P. M. (1987) 'TRANSFER RNA MODIFICATION', *Ann. Rev. Biochem*, 56, pp. 263-87.
- Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortés-Garrido, R., Gkatza, N., Dietmann, S. and Frye, M. (2016) 'Stem cell function and stress response are controlled by protein synthesis', *Nature*. Nature Publishing Group, 534(7607), pp. 335-340.
- Blanco, S., Dietmann, S., Flores, J. V., Hussain, S., Kutter, C., Humphreys, P., Lukk, M., Lombard, P., Treps, L., Popis, M., Kellner, S., Holter, S. M., Garrett, L., Wurst, W., Becker, L., Klopstock, T., Fuchs, H., Gailus-Durner, V., Hrabe de Angelis, M., Karadottir, R. T., Helm, M., Ule, J., Gleeson, J. G., Odom, D. T. and Frye, M., (2014) 'Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders', *The EMBO Journal*. EMBO Press, 33(18), pp. 2020-2039.

- Blanco, S., Kurowski, A., Nichols, J., Watt, F. M., Benitah, S. A. and Frye, M. (2011) 'The RNA-Methyltransferase Miso (NSun2) Poises Epidermal Stem Cells to Differentiate', *PLoS Genetics*. 7(12), p. e1002403.
- Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T. and Weber, M. (2010) 'Targets and dynamics of promoter DNA methylation during early mouse development', *Nature Genetics*. Nature Publishing Group, 42(12), pp. 1093–1100.
- Boroviak, T. and Nichols, J. (2014) 'The birth of embryonic pluripotency.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. The Royal Society, 369(1657).
- Boroviak, T., Loos, R., Bertone, P., Smith, A. and Nichols, J. (2014) 'The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification', *Nature Cell Biology*. Nature Publishing Group, 16(6), pp. 513–525.
- Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A. and Bertone, P. (2015) 'Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis', *Developmental Cell*. Cell Press, 35(3), pp. 366–382.
- Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E. (1984) 'Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines', *Nature*. Nature Publishing Group, 309(5965), pp. 255–256.
- Bujnicki, J. M., Feder, M., Ayres, C. L. and Redman, K. L. (2004) 'Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases.', *Nucleic acids research*, 32(8), pp. 2453–63.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., Davey, N. E., Humphreys, D. T., Preiss, T., Steinmetz, L. M., Krijgsveld, J. and Hentze, M. W. (2012) 'Insights into RNA biology from an atlas of mammalian mRNA-binding proteins.', *Cell*, 149(6), pp. 1393–406.
- Chazaud, C., Yamanaka, Y., Pawson, T. and Rossant, J. (2006) 'Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway', *Developmental Cell*. Cell Press, 10(5), pp. 615–624.
- Chen, T., Hao, Y.-J., Zhang, Y., Li, M.-M., Wang, M., Han, W., Wu, Y., Lv, Y., Hao, J., Wang, L., Li, A., Yang, Y., Jin, K.-X., Zhao, X., Li, Y., Ping, X.-L., Lai, W.-Y., Wu, L.-G., Jiang, G., *et al.* (2015) 'm6A RNA Methylation Is Regulated by MicroRNAs and Promotes Reprogramming to Pluripotency', *Cell Stem Cell*, 16(3), pp. 289–301.
- Chen, Y., Sierzputowska-Gracz, H., Guenther, R., Everett, K. and Agris, P. F. (1993) '5-Methylcytidine is required for cooperative binding of Mg<sup>2+</sup> and a conformational transition at the anticodon stem-loop of yeast phenylalanine tRNA.', *Biochemistry*, 32(38), pp. 10249–53.

- Cheng, J. X., Chen, L., Li, Y., Cloe, A., Yue, M., Wei, J., Watanabe, K. A., Shammo, J. M., Anastasi, J., Shen, Q. J., Larson, R. A., He, C., Le Beau, M. M. and Vardiman, J. W. (2018) 'RNA cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia', *Nature Communications*. Nature Publishing Group, 9(1), p. 1163.
- Choi, S. Y., Choi, M. J., Cho, M. Y. and Park, Y. J. (2019) 'Alkbh5, a RNA Demethylase, Is Involved in Fine-tuning of Cell Differentiation.', *Current developments in nutrition*. American Society for Nutrition, 3(Suppl 1).
- Choi, J., Huebner, A. J., Clement, K., Walsh, R. M., Savol, A., Lin, K., Gu, H., Di Stefano, B., Brumbaugh, J., Kim, S.-Y., Sharif, J., Rose, C. M., Mohammad, A., Odajima, J., Charron, J., Shioda, T., Gnirke, A., Gygi, S., Koseki, H., Sadreyev, R. I., Xiao, A., Meissner, A. and Hochedlinger, K. (2017) 'Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells', *Nature*. Nature Research, 548(7666), p. 219.
- Chrysanthou, S., Senner, C. E., Woods, L., Fineberg, E., Okkenhaug, H., Burge, S., Perez-Garcia, V. and Hemberger, M. (2018) 'A Critical Role of TET1/2 Proteins in Cell-Cycle Progression of Trophoblast Stem Cells', *Stem Cell Reports*. Cell Press, 10(4), pp. 1355-1368.
- Clancy, M. J., Shambaugh, M. E., Timpte C. S. and 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 Research*, 30(20), pp. 4509-4518.
- Cockburn, K. and Rossant, J. (2010) 'Making the blastocyst: lessons from the mouse.', *The Journal of clinical investigation*. American Society for Clinical Investigation, 120(4), pp. 995-1003.
- Cui, W., Pizzollo, J., Han, Z., Marcho, C., Zhang, K. and Mager, J. (2016) 'Nop2 is required for mammalian preimplantation development.', *Molecular reproduction and development*, 83(2), pp. 124-31.
- Cui, X., Liang, Z., Shen, L., Zhang, Q., Bao, S., Geng, Y., Zhang, B., Leo, V., Vardy, L. A., Lu, T., Gu, X. and Yu, H. (2017) '5-Methylcytosine RNA Methylation in *Arabidopsis Thaliana*', *Molecular Plant*. Cell Press, 10(11), pp. 1387-1399.
- Delatte, B., Wang, F., Ngoc, L. V., Collignon, E., Bonvin, E., Deplus, R., Calonne, E., Hassabi, B., Putmans, P., Awe, S., Wetzel, C., Kreher, J., Soin, R., Creppe, C., Limbach, P. A., Gueydan, C., Kruys, V., Brehm, A., Minakhina, S., Defrance, M., Steward, R. and Fuks, F. (2016) 'RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 351(6270), pp. 282-5.
- Dietrich, J.-E. and Hiiragi, T. (2007) 'Development', *Development*. The Company of Biologists Ltd, p. 815.

- 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. and Rechavi, G. (2012) 'Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 485(7397), pp. 201-6.
- Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J. and Wu, L. (2016) 'YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex', *Nature Communications*. Nature Publishing Group, 7(1), p. 12626.
- Duncan, B. K. and Miller, J. H. (1980) 'Mutagenic deamination of cytosine residues in DNA', *Nature*. Nature Publishing Group, 287(5782), pp. 560-561.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P. and Ehresmann, B. (1987) 'Probing the structure of RNAs in solution', *Nucleic Acids Research*. Narnia, 15(22), pp. 9109-9128.
- Fawcett, K. A. and Barroso, I. (2010) 'The genetics of obesity: FTO leads the way', *Trends in Genetics*. Elsevier Current Trends, 26(6), pp. 266-274.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1990) 'Electrospray ionization-principles and practice', *Mass Spectrometry Reviews*. John Wiley & Sons, Ltd, 9(1), pp. 37-70.
- Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A., Marques, C. J., Andrews, S. and Reik, W. (2011) 'Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation', *Nature*. Nature Publishing Group, 473(7347), pp. 398-402.
- Ficz, G., Hore, T. A., Santos, F., Lee, H. J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.-L., Walter, J., Cook, S. J., Andrews, S., Branco, M. R. and Reik, W. (2013) 'FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency', *Cell Stem Cell*. Elsevier, 13(3), pp. 351-359.
- Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R. and Bickmore, W. A. (2008) 'Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells', *PLoS Genetics*. Edited by W. Reik. Public Library of Science, 4(3), p. e1000039.
- Fischer, J., Koch, L., Emmerling, C., Vierkotten, J., Peters, T., Brüning, J. C. and Rüther, U. (2009) 'Inactivation of the Fto gene protects from obesity', *Nature*. Nature Publishing Group, 458(7240), pp. 894-898.
- Fléchon, J.-E. and Kopecký, V. (1998) 'The nature of the "nucleolus precursor body" in early preimplantation embryos: a review of fine-structure cytochemical, immunocytochemical and autoradiographic data related to nucleolar function', *Zygote*. Cambridge University Press, 6(2), pp. 183-191.

- FLEISSNER, E. and BOREK, E. (1962) 'A new enzyme of RNA synthesis: RNA methylase.', *Proceedings of the National Academy of Sciences of the United States of America*, 48, pp. 1199–203.
- Fonagy, A., Swiderski, C., Dunn, M. and Freeman, J. W. (1992) 'Antisense-mediated Specific Inhibition of P120 Protein Expression Prevents G1- to S-Phase Transition', *Cancer Res.*, 52(19), pp. 5250–5256.
- Fonagy, A., Swiderski, C., Wilson, A., Bolton, W., Kenyon, N. and Freeman, J. W. (1993) 'Cell cycle regulated expression of nucleolar antigen P120 in normal and transformed human fibroblasts.', *Journal of cellular physiology*, 154(1), pp. 16–27.
- Freeman, J. W., Busch, R. K., Gyorkey, F., Gyorkey, P., Ross, B. E. and Busch, H. (1988) 'Identification and characterization of a human proliferation-associated nucleolar antigen with a molecular weight of 120,000 expressed in early G1 phase.', *Cancer research*, 48(5), pp. 1244–51.
- Frye, M. and Watt, F. M. (2006) 'The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors.', *Current biology : CB*, 16(10), pp. 971–81.
- Fu, L., Guerrero, I., Zhong, N., Amato, N. J., Liu, Y., Liu, S., Cai, Q., Ji, D., Jin, S.-G., Niedernhofer, L. J., Pfeifer, G. P., Xu, G.-L. and Wang, Y. (2014) 'Tet-mediated formation of 5-hydroxymethylcytosine in RNA.', *Journal of the American Chemical Society*. American Chemical Society, 136(33), pp. 11582–5.
- Fujiwara, Y., Komiya, T., Kawabata, H., Satoh, M., Fujimoto, H., Furusawa, M. and Noctor, T. (1994) 'Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila vasa and its specific expression in germ cell lineage', *Developmental Biology*, 91, pp. 12258–12262.
- Furuichi, Y., LaFiandra, A. and Shatkin, A. J. (1977) '5'-Terminal structure and mRNA stability.', *Nature*, 266(5599), pp. 235–9.
- Geier, G. E. and Modrich, P. (1979) 'Recognition sequence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease.', *The Journal of biological chemistry*, 254(4), pp. 1408–13.
- Gerbe, F., Cox, B., Rossant, J. and Chazaud, C. (2008) 'Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst', *Developmental Biology*. Academic Press, 313(2), pp. 594–602.
- 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., *et al.* (2015) 'Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation.',

*Science (New York, N.Y.)*. American Association for the Advancement of Science, 347(6225), pp. 1002–6.

Ginsburg, M., Snow, M. H. and McLaren, A. (1990) ‘Primordial germ cells in the mouse embryo during gastrulation’, *Development*, 110(2).

Goldberg, A. D., Allis, C. D. and Bernstein, E. (2007) ‘Epigenetics: A Landscape Takes Shape’, *Cell*. Cell Press, 128(4), pp. 635–638.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C.-L., Zhang, X., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006) ‘Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2.’, *Science (New York, N.Y.)*, 311(5759), pp. 395–8.

Gowher, H., Leismann, O. and Jeltsch, A. (2000) ‘DNA of *Drosophila melanogaster* contains 5-methylcytosine’, *The EMBO Journal*. EMBO Press, 19(24), pp. 6918–6923.

Habibi, E., Brinkman, A. B., Arand, J., Kroeze, L. I., Kerstens, H. H. D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N. C., Benedetti, R., Altucci, L., Jansen, J. H., Walter, J., Gut, I. G., Marks, H. and Stunnenberg, H. G. (2013) ‘Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells’, *Cell Stem Cell*. Elsevier, 13(3), pp. 360–369.

Hackett, J. A., Sengupta, R., Zylitz, J. J., Murakami, K., Lee, C., Down, T. A. and Surani, M. A. (2013) ‘Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine.’, *Science (New York, N.Y.)*. American Association for the Advancement of Science, 339(6118), pp. 448–52.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M. and Tuschl, T. (2010) ‘Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP’, *Cell*. Cell Press, 141(1), pp. 129–141.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U. C., Cesari, F., Lee, C., Almouzni, G., Schneider, R. and Surani, M. A. (2008) ‘Chromatin dynamics during epigenetic reprogramming in the mouse germ line’, *Nature*, 452(7189), pp. 877–881.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J. and Surani, M. A. (2002) ‘Epigenetic reprogramming in mouse primordial germ cells’, *Mechanisms of Development*. Elsevier, 117(1–2), pp. 15–23.

Hajkova, P., Jeffries, S. J., Lee, C., Miller, N., Jackson, S. P. and Surani, M. A. (2010) ‘Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway.’, *Science (New York, N.Y.)*. American Association for the Advancement of Science, 329(5987), pp. 78–82.

Harvy, D. (2014) Contributions to Band Broadening in Chromatography. [online] Available at: <https://community.asdlib.org/imageandvideoexchange/forum/2013/08/01/contributions-to-band-broadening-in-chromatography> [Accessed 2 jul. 2019]

Haupt, A., Thamer, C., Staiger, H., Tschritter, O., Kirchhoff, K., Machicao, F., Häring, H.-U., Stefan, N. and Fritsche, A. (2008) 'Variation in the FTO Gene Influences Food Intake but not Energy Expenditure', *Experimental and Clinical Endocrinology & Diabetes*. © J. A. Barth Verlag in Georg Thieme Verlag KG Stuttgart · New York, 117(04), pp. 194–197.

Hayashi, K., Lopes, S. M. C. de S., Tang, F. and Surani, M. A. (2008) 'Dynamic Equilibrium and Heterogeneity of Mouse Pluripotent Stem Cells with Distinct Functional and Epigenetic States', *Cell Stem Cell*. Cell Press, 3(4), pp. 391–401.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. and Saitou, M. (2011) 'Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells', *Cell*. Cell Press, 146(4), pp. 519–532.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.-X., Zhang, K., He, C. and Xu, G.-L. (2011) 'Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA.', *Science (New York, N.Y.)*, 333(6047), pp. 1303–7.

Heard, E. and Martienssen, R. A. (2014) 'Transgenerational epigenetic inheritance: myths and mechanisms.', *Cell*. NIH Public Access, 157(1), pp. 95–109.

Hernandez-Verdun, D., Roussel, P., Thiry, M., Sirri, V. and Lafontaine, D. L. J. (2010) 'The nucleolus: structure/function relationship in RNA metabolism', *Wiley Interdisciplinary Reviews: RNA*. John Wiley & Sons, Ltd, 1(3), pp. 415–431.

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., Rütther, U., Jaffrey, S. R., Kloppenburg, P. and Brünig, J. C. (2013) 'The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry', *Nature Neuroscience*. Nature Publishing Group, 16(8), pp. 1042–1048.

Hill, P. W. S., Amouroux, R. and Hajkova, P. (2014) 'DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: An emerging complex story', *Genomics*, 104(5), pp. 324–333.

Hill, P. W. S., Leitch, H. G., Requena, C. E., Sun, Z., Amouroux, R., Roman-Trufero, M., Borkowska, M., Terragni, J., Vaisvila, R., Linnett, S., Bagci, H., Dharmalingham, G., Haberer, V., Lenhard, B., Zheng, Y., Pradhan, S. and Hajkova, P. (2018) 'Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte', *Nature*, 555(7696), pp. 392–396.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. and Sasaki, H. (2008) 'Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA

methylation imprints during preimplantation development.’, *Genes & development*. Cold Spring Harbor Laboratory Press, 22(12), pp. 1607–16.

Hodge, R. (1998) ‘Preparation of RNA Dot-Blots’, in *RNA Isolation and Characterization Protocols*. Totowa, NJ: Humana Press, pp. 73–75.

Hoernes, T. P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., Lindner, H., Hüttenhofer, A. and Erlacher, M. D. (2016) ‘Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code’, *Nucleic Acids Research*. Oxford University Press, 44(2), pp. 852–862.

Hong, B., Brockenbrough, J. S., Wu, P. and Aris, J. P. (1997) ‘Nop2p is required for pre-rRNA processing and 60S ribosome subunit synthesis in yeast.’, *Molecular and Cellular Biology*, 17(1), pp. 378–388.

Hong, J., Lee, J. H. and Chung, I. K. (2016) ‘Telomerase activates transcription of cyclin D1 gene through an interaction with NOL1’, *Journal of Cell Science*. The Company of Biologists Ltd, 129(8), pp. 1566–1579.

Howlett, S. K. and Reik, W. (1991) ‘Methylation levels of maternal and paternal genomes during preimplantation development’, *Development*, 113(1).

Huang, T., Chen, W., Liu, J., Gu, N. and Zhang, R. (2019) ‘Genome-wide identification of mRNA 5-methylcytosine in mammals’, *Nature Structural & Molecular Biology*. Nature Publishing Group, 26(5), pp. 380–388.

Huang, W., Lan, M.-D., Qi, C.-B., Zheng, S.-J., Wei, S.-Z., Yuan, B.-F. and Feng, Y.-Q. (2016) ‘Formation and determination of the oxidation products of 5-methylcytosine in RNA’, *Chemical Science*. Royal Society of Chemistry, 7(8), pp. 5495–5502.

Huang, Y., Osorno, R., Tsakiridis, A. and Wilson, V. (2012) ‘In Vivo Differentiation Potential of Epiblast Stem Cells Revealed by Chimeric Embryo Formation’, *Cell Reports. Cell Press*, 2(6), pp. 1571–1578.

Huber, S. M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E. A. and Balasubramanian, S. (2015) ‘Formation and abundance of 5-hydroxymethylcytosine in RNA.’, *Chembiochem : a European journal of chemical biology*, 16(5), pp. 752–5.

Hussain, S., Benavente, S. B., Nascimento, E., Dragoni, I., Kurowski, A., Gillich, A., Humphreys, P. and Frye, M. (2009) ‘The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability.’, *The Journal of cell biology*, 186(1), pp. 27–40.

Hussain, S., Sajini, A. A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., Paramor, M., Gleeson, J. G., Odom, D. T., Ule, J. and Frye, M. (2013) ‘NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs.’, *Cell reports*, 4(2), pp. 255–61.

Inoue, A. and Aoki, F. (2010) 'Role of the nucleoplasmin 2 C-terminal domain in the formation of nucleolus-like bodies in mouse oocytes', *The FASEB Journal*. Federation of American Societies for Experimental Biology, 24(2), pp. 485-494.

Irie, N., Weinberger, L., Tang, W. W. C., Kobayashi, T., Viukov, S., Manor, Y. S., Dietmann, S., Hanna, J. H. and Surani, M. A. (2015) 'SOX17 is a critical specifier of human primordial germ cell fate.', *Cell*. Elsevier, 160(1-2), pp. 253-68.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. and Zhang, Y. (2011) 'Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.', *Science (New York, N.Y.)*, 333(6047), pp. 1300-3.

Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P. N., Monahan, J., Carrieri, C., Enright, A. J. and O'Carroll, D. (2017) 'The RNA m6A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence', *Molecular Cell*. Cell Press, 67(6), pp. 1059-1067.e4.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G. and He, C. (2011) 'N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO.', *Nature chemical biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 7(12), pp. 885-7.

Jurkowski, T. P. and Jeltsch, A. (2011) 'On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2.', *PloS one*. Public Library of Science, 6(11), p. e28104.

Jurkowski, T. P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G. and Jeltsch, A. (2008) 'Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism.', *RNA (New York, N.Y.)*, 14(8), pp. 1663-70.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. and Razin, A. (1992) 'Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line.', *Genes & Development*. Cold Spring Harbor Laboratory Press, 6(5), pp. 705-714.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. and Saitou, M. (2012) 'Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice', *The EMBO Journal*. John Wiley & Sons, Ltd, 32(3), pp. 340-353.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H. J., Nett, I., Marks, H., Walker, R., Stunnenberg, H. G., Lilley, K. S., Nichols, J., Reik, W., Bertone, P. and Smith, A. (2017) 'Tracking the embryonic stem cell transition from ground state pluripotency', *Development*. Oxford University Press for The Company of Biologists Limited, 144(7), pp. 1221-1234.

Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., Schultz, R. M. and Wang, P. J. (2018) 'Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development', *PLOS Genetics*. Edited by W. Yan, 14(5), p. e1007412.

- 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. and Darnell, R. B. (2015) 'A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation.', *Genes & development*. Cold Spring Harbor Laboratory Press, 29(19), pp. 2037–53.
- Keshet, I., Yisraeli, J. and Cedar, H. (1985) 'Effect of regional DNA methylation on gene expression.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 82(9), pp. 2560–4.
- Khoddami, V. and Cairns, B. R. (2013) 'Identification of direct targets and modified bases of RNA cytosine methyltransferases.', *Nature biotechnology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 31(5), pp. 458–64.
- Kierzek, E. and Kierzek, R. (2003) 'The thermodynamic stability of RNA duplexes and hairpins containing N6-alkyladenosines and 2-methylthio-N6-alkyladenosines', *Nucleic Acids Research*. Narnia, 31(15), pp. 4472–4480.
- Kim, H., Jang, H., Kim, T. W., Kang, B.-H., Lee, S. E., Jeon, Y. K., Chung, D. H., Choi, J., Shin, J., Cho, E.-J. and Youn, H.-D. (2015) 'Core Pluripotency Factors Directly Regulate Metabolism in Embryonic Stem Cell to Maintain Pluripotency', *STEM CELLS*. John Wiley & Sons, Ltd, 33(9), pp. 2699–2711.
- Kim, H., Wu, J., Ye, S., Tai, C.-I., Zhou, X., Yan, H., Li, P., Pera, M. and Ying, Q.-L. (2013) 'Modulation of  $\beta$ -catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal', *Nature Communications*. Nature Research, 4.
- Knuckles, P., Carl, S. H., Musheev, M., Niehrs, C., Wenger, A. and Bühler, M. (2017) 'RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding', *Nature Structural & Molecular Biology*. Nature Publishing Group, 24(7), pp. 561–569.
- Koh, K. P., Yabuuchi, A., Rao, S., Huang, Y., Cuniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C. A., Mostoslavsky, G., Lahesmaa, R., Orkin, S. H., Rodig, S. J., Daley, G. Q. and Rao, A. (2011) 'Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells.', *Cell stem cell*, 8(2), pp. 200–13.
- Kondrashov, A. S. (2003) 'Direct estimates of human per nucleotide mutation rates at 20 loci causing mendelian diseases', *Human Mutation*. John Wiley & Sons, Ltd, 21(1), pp. 12–27.
- Kosi, N., Alić, I., Kolačević, M., Vrsaljko, N., Jovanov Milošević, N., Sobol, M., Philimonenko, A., Hozák, P., Gajović, S., Pochet, R. and Mitrečić, D. (2015) 'Nop2 is expressed during proliferation of neural stem cells and in adult mouse and human brain.', *Brain research*, 1597, pp. 65–76.

- Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E. and Allis, C. D. (1998) 'Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo.', *Genes & development*. Cold Spring Harbor Laboratory Press, 12(5), pp. 627-39.
- Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K. and Saitou, M. (2008) 'Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice.', *Genes & development*. Cold Spring Harbor Laboratory Press, 22(12), pp. 1617-35.
- Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. and Reik, W. (2003) 'Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse', *genesis*. John Wiley & Sons, Ltd, 35(2), pp. 88-93.
- Larijani, M., Frieder, D., Sonbuchner, T. M., Bransteitter, R., Goodman, M. F., Bouhassira, E. E., Scharff, M. D. and Martin, A. (2005) 'Methylation protects cytidines from AID-mediated deamination', *Molecular Immunology*. Pergamon, 42(5), pp. 599-604.
- Lawson, K. A., Dunn, N. R., Roelen, B. A. J., Zeinstra, L. M., Davis, A. M., Wright, C. V. E., Korving, J. P. W. F. M. and Hogan, B. L. M. (1999) 'Bmp4 is required for the generation of primordial germ cells in the mouse embryo', *Genes & Development*. Cold Spring Harbor Laboratory Press, 13(4), pp. 424-436.
- Leitch, H. G., McEwen, K. R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J. G., Smith, A., Surani, M. A. and Hajkova, P. (2013) 'Naive pluripotency is associated with global DNA hypomethylation.', *Nature structural & molecular biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 20(3), pp. 311-6.
- Lewinska, A., Wnuk, M., Grabowska, W., Zabek, T., Semik, E., Sikora, E. and Bielak-Zmijewska, A. (2015) 'Curcumin induces oxidation-dependent cell cycle arrest mediated by SIRT7 inhibition of rDNA transcription in human aortic smooth muscle cells', *Toxicology Letters*. Elsevier, 233(3), pp. 227-238.
- Li, H. H., Roy, M., Kuscuoglu, U., Spencer, C. M., Halm, B., Harrison, K. C., Bayle, J. H., Splendore, A., Ding, F., Meltzer, L. A., Wright, E., Paylor, R., Deisseroth, K. and Francke, U. (2009) 'Induced chromosome deletions cause hypersociability and other features of Williams-Beuren syndrome in mice', *EMBO Molecular Medicine*, 1(1), pp. 50-65.
- Li, L., Zheng, P. and Dean, J. (2010) 'Maternal control of early mouse development', *Development*. Oxford University Press for The Company of Biologists Limited, 137(6), pp. 859-870.
- Limbach, P. A., Crain, P. F. and McCloskey, J. A. (1994) 'Summary: the modified nucleosides of RNA.', *Nucleic acids research*, 22(12), pp. 2183-96.

Lister, R., Pelizzola, M., Downen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q.-M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B. and Ecker, J. R. (2009) 'Human DNA methylomes at base resolution show widespread epigenomic differences', *Nature*. Nature Publishing Group, 462(7271), pp. 315-322.

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W. and He, C. (2014) 'A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation.', *Nature chemical biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 10(2), pp. 93-5.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M. and Pan, T. (2015) 'N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 518(7540), pp. 560-564.

Lu, P., Vogel, C., Wang, R., Yao, X. and Marcotte, E. M. (2007) 'Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation', *Nature Biotechnology*. Nature Publishing Group, 25(1), pp. 117-124.

Lyko, F., Ramsahoye, B. H. and Jaenisch, R. (2000) 'DNA methylation in *Drosophila melanogaster*', *Nature*. Nature Publishing Group, 408(6812), pp. 538-540.

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. and Grosjean, H. (2013) 'MODOMICS: a database of RNA modification pathways-2013 update.', *Nucleic acids research*, 41(Database issue), pp. D262-7.

Magnúsdóttir, E., Dietmann, S., Murakami, K., Günesdogan, U., Tang, F., Bao, S., Diamanti, E., Lao, K., Gottgens, B. and Azim Surani, M. (2013) 'A tripartite transcription factor network regulates primordial germ cell specification in mice', *Nature Cell Biology*. Nature Publishing Group, 15(8), pp. 905-915.

Mantsoki, A., Devailly, G. and Joshi, A. (2015) 'CpG island erosion, polycomb occupancy and sequence motif enrichment at bivalent promoters in mammalian embryonic stem cells', *Scientific Reports*. Nature Publishing Group, 5(1), p. 16791.

Marinus, M. G. and Morris, N. R. (1973) 'Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12.', *Journal of bacteriology*, 114(3), pp. 1143-50.

Martinez, F. J., Lee, J. H., Lee, J. E., Blanco, S., Nickerson, E., Gabriel, S., Frye, M., Al-Gazali, L. and Gleeson, J. G. (2012) 'Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome.', *Journal of medical genetics*, 49(6), pp. 380-5.

Masaki, H., Kato-Itoh, M., Takahashi, Y., Umino, A., Sato, H., Ito, K., Yanagida, A., Nishimura, T., Yamaguchi, T., Hirabayashi, M., Era, T., Loh, K. M., Wu, S. M., Weissman, I. L. and

Nakauchi, H. (2016) 'Inhibition of Apoptosis Overcomes Stage-Related Compatibility Barriers to Chimera Formation in Mouse Embryos', *Cell Stem Cell*. Cell Press, 19(5), pp. 587–592.

May, M. S. and Hattman, S. (1975) 'Analysis of bacteriophage deoxyribonucleic acid sequences methylated by host- and R-factor-controlled enzymes.', *Journal of bacteriology*, 123(2), pp. 768–70.

Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000) 'Demethylation of the zygotic paternal genome', *Nature*. Nature Publishing Group, 403(6769), pp. 501–502.

Mendel, M., Chen, K.-M., Homolka, D., Gos, P., Pandey, R. R., McCarthy, A. A. and Pillai, R. S. (2018) 'Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development', *Molecular Cell*. Cell Press, 71(6), pp. 986-1000.e11.

Merkenschlager, M. and Nora, E. P. (2016) 'CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation', *Annual Review of Genomics and Human Genetics*. Annual Reviews , 17(1), pp. 17–43.

Messerschmidt, D. M., Knowles, B. B. and Solter, D. (2014) 'DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos.', *Genes & development*, 28(8), pp. 812–28.

Metodiev, M. D., Spähr, H., Loguercio Polosa, P., Meharg, C., Becker, C., Altmueller, J., Habermann, B., Larsson, N.-G. and Ruzzenente, B. (2014) 'NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly.', *PLoS genetics*. Public Library of Science, 10(2), p. e1004110.

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E. and Jaffrey, S. R. (2012) 'Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and near Stop Codons', *Cell*. Cell Press, 149(7), pp. 1635–1646.

Miao, Z., Xin, N., Wei, B., Hua, X., Zhang, G., Leng, C., Zhao, C., Wu, D., Li, J., Ge, W., Sun, M. and Xu, X. (2016) '5-hydroxymethylcytosine is detected in RNA from mouse brain tissues', *Brain Research*, 1642, pp. 546–552.

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., *et al.* (2007) 'Genome-wide maps of chromatin state in pluripotent and lineage-committed cells', *Nature*. Nature Publishing Group, 448(7153), pp. 553–560.

Miller, P. E. and Denton, M. B. (1986) 'The quadrupole mass filter: Basic operating concepts', *Journal of Chemical Education*, 63(7), p. 617.

Mishima, E., Jinno, D., Akiyama, Y., Itoh, K., Nankumo, S., Shima, H., Kikuchi, K., Takeuchi, Yoichi, Elkordy, A., Suzuki, T., Niizuma, K., Ito, S., Tomioka, Y., Abe, T., Cantara, W., Crain,

P., Rozenski, J., McCloskey, J., Harris, K., *et al.* (2015) 'Immuno-Northern Blotting: Detection of RNA Modifications by Using Antibodies against Modified Nucleosides', *PLOS ONE*. Edited by T. Preiss. Public Library of Science, 10(11), p. e0143756.

Monk, M., Boubelik, M. and Lehnert, S. (1987) 'Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development', *Development*, 99(3).

Mora, J. F. de la, Van Berkel, G. J., Enke, C. G., Cole, R. B., Martinez-Sanchez, M. and Fenn, J. B. (2000) 'Electrochemical processes in electrospray ionization mass spectrometry', *Journal of Mass Spectrometry*. John Wiley & Sons, Ltd, 35(8), pp. 939-952.

Morgan, H. D., Dean, W., Coker, H. A., Reik, W. and Petersen-Mahrt, S. K. (2004) 'Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 279(50), pp. 52353-60.

Motorin, Y., Lyko, F. and Helm, M. (2010) '5-methylcytosine in RNA: detection, enzymatic formation and biological functions.', *Nucleic acids research*, 38(5), pp. 1415-30.

Murakami, Y., Spriggs, R. V., Nakamura, H. and Jones, S. (2010) 'PiRaNhA: a server for the computational prediction of RNA-binding residues in protein sequences', *Nucleic Acids Research*. Oxford Academic, 38(Web Server), pp. W412-W416.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000) 'Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme', *Cell*. Cell Press, 102(5), pp. 553-563.

Nabel, C. S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H. L., Stivers, J. T., Zhang, Y. and Kohli, R. M. (2012) 'AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation', *Nature Chemical Biology*. Nature Publishing Group, 8(9), pp. 751-758.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T. and Saitou, M. (2016) 'A developmental coordinate of pluripotency among mice, monkeys and humans', *Nature*. Nature Publishing Group, 537(7618), pp. 57-62.

Nashun, B., Hill, P. W. S., Smallwood, S. A., Dharmalingam, G., Amouroux, R., Clark, S. J., Sharma, V., Ndjetehe, E., Pelczar, P., Festenstein, R. J., Kelsey, G. and Hajkova, P. (2015) 'Continuous Histone Replacement by Hira Is Essential for Normal Transcriptional Regulation and De Novo DNA Methylation during Mouse Oogenesis', *Molecular Cell*. Cell Press, 60(4), pp. 611-625.

Nashun, B., Yukawa, M., Liu, H., Akiyama, T. and Aoki, F. (2010) 'Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice.',

*Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 137(22), pp. 3785–94.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R. and Rossant, J. (2005) 'Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation', *Cell*. Cell Press, 123(5), pp. 917–929.

O'Connor, M. D., Kardel, M. D., Iosfina, I., Youssef, D., Lu, M., Li, M. M., Vercauteren, S., Nagy, A. and Eaves, C. J. (2008) 'Alkaline Phosphatase-Positive Colony Formation Is a Sensitive, Specific, and Quantitative Indicator of Undifferentiated Human Embryonic Stem Cells', *Stem Cells*. John Wiley & Sons, Ltd, 26(5), pp. 1109–1116.

O'Farrell, P. H., Stumpff, J. and Tin Su, T. (2004) 'Embryonic Cleavage Cycles: How Is a Mouse Like a Fly?', *Current Biology*. Cell Press, 14(1), pp. R35–R45.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S. C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., Saitou, M. and Surani, M. A. (2005) 'Blimp1 is a critical determinant of the germ cell lineage in mice', *Nature*. Nature Publishing Group, 436(7048), pp. 207–213.

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T. and Saitou, M. (2009) 'A Signaling Principle for the Specification of the Germ Cell Lineage in Mice', *Cell*. Cell Press, 137(3), pp. 571–584.

Okamoto, M., Hirata, S., Sato, S., Koga, S., Fujii, M., Qi, G., Ogawa, I., Takata, T., Shimamoto, F. and Tatsuka, M. (2012) 'Frequent increased gene copy number and high protein expression of tRNA (cytosine-5-)-methyltransferase (NSUN2) in human cancers.', *DNA and cell biology*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 31(5), pp. 660–71.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999) 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development', *Cell*, 99(3), pp. 247–257.

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P. J., Chambers, I. and Wilson, V. (2012) 'The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 139(13), pp. 2288–98.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. and Walter, J. (2000) 'Active demethylation of the paternal genome in the mouse zygote', *Current Biology*. Cell Press, 10(8), pp. 475–478.

Patil, D. P., Chen, C.-K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M. and Jaffrey, S. R. (2016) 'm6A RNA methylation promotes XIST-mediated transcriptional repression', *Nature*. Nature Research, 537(7620), pp. 369–373.

- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P. and Conrad, N. K. (2017) 'The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention', *Cell*. Cell Press, 169(5), pp. 824-835.e14.
- Peng, S., Chen, L.-L., Lei, X.-X., Yang, L., Lin, H., Carmichael, G. G. and Huang, Y. (2011) 'Genome-Wide Studies Reveal That Lin28 Enhances the Translation of Genes Important for Growth and Survival of Human Embryonic Stem Cells', *STEM CELLS*. John Wiley & Sons, Ltd, 29(3), pp. 496-504. doi: 10.1002/stem.591.
- Perry, R. P., Kelley, D. E., Friderici, K. and Rottman, F. (1975) 'The methylated constituents of L cell messenger RNA: Evidence for an unusual cluster at the 5' terminus', *Cell*, 4(4), pp. 387-394.
- Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M. C. and Reuter, G. (2009) 'Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2.', *Nature genetics*. Nature Publishing Group, 41(6), pp. 696-702.
- Polesskaya, A., Cuvellier, S., Naguibneva, I., Duquet, A., Moss, E. G. and Harel-Bellan, A. (2007) 'Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency.', *Genes & development*. Cold Spring Harbor Laboratory Press, 21(9), pp. 1125-38. doi: 10.1101/gad.415007.
- Poole, A., Penny, D. and Sjöberg, B.-M. (2001) 'Confounded cytosine! Tinkering and the evolution of DNA', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 2(2), pp. 147-151.
- Popp, C., Dean, W., Feng, S., Cokus, S. J., Andrews, S., Pellegrini, M., Jacobsen, S. E. and Reik, W. (2010) 'Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency', *Nature*. Nature Publishing Group, 463(7284), pp. 1101-1105.
- Prioleau, M. N., Huet, J., Sentenac, A. and Méchali, M. (1994) 'Competition between chromatin and transcription complex assembly regulates gene expression during early development.', *Cell*. Elsevier, 77(3), pp. 439-49.
- Radford, E. J., Ito, M., Shi, H., Corish, J. A., Yamazawa, K., Isganaitis, E., Seisenberger, S., Hore, T. A., Reik, W., Erkek, S., Peters, A. H. F. M., Patti, M.-E. and Ferguson-Smith, A. C. (2014) 'In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism.', *Science (New York, N.Y.)*. NIH Public Access, 345(6198), p. 1255903.
- Rai, K., Chidester, S., Zavala, C. V., Manos, E. J., James, S. R., Karpf, A. R., Jones, D. A. and Cairns, B. R. (2007) 'Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish', *Genes & Development*, 21(3), pp. 261-266.

- Rajendram, R., Ferreira, J. C., Grafodatskaya, D., Choufani, S., Chiang, T., Pu, S., Butcher, D. T., Wodak, S. J. and Weksberg, R. (2011) 'Assessment of methylation level prediction accuracy in methyl-DNA immunoprecipitation and sodium bisulfite based microarray platforms', *Epigenetics*. Taylor & Francis, 6(4), pp. 410–415.
- Ralston, A. and Rossant, J. (2008) 'Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo', *Developmental Biology*. Academic Press, 313(2), pp. 614–629.
- Ramakrishnan, S. R., Vogel, C., Prince, J. T., Wang, R., Li, Z., Penalva, L. O., Myers, M., Marcotte, E. M. and Miranker, D. P. (2009) 'Integrating shotgun proteomics and mRNA expression data to improve protein identification', *Bioinformatics*. Narnia, 25(11), pp. 1397–1403.
- Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I. and Cuzin, F. (2006) 'RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse', *Nature*. Nature Publishing Group, 441(7092), pp. 469–474.
- Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Lagelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A. G., Brousse, N., Muramatsu, M., Notarangelo, L. D., Kinoshita, K., Honjo, T., *et al.* (2000) 'Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2)', *Cell*. Cell Press, 102(5), pp. 565–575.
- Robertson, E. J. (2014) 'Dose-dependent Nodal/Smad signals pattern the early mouse embryo', *Seminars in Cell & Developmental Biology*. Academic Press, 32, pp. 73–79.
- ROE, B., MICHAEL, M. and DUDOCK, B. (1973) 'Function of N2 Methylguanine in Phenylalanine Transfer RNA', *Nature*, *Published online: 05 December 1973*; / *doi:10.1038/10.1038/newbio246135a0*. Nature Publishing Group, 246(153), p. 135.
- Rošić, S., Amouroux, R., Requena, C. E., Gomes, A., Emperle, M., Beltran, T., Rane, J. K., Linnett, S., Selkirk, M. E., Schiffer, P. H., Bancroft, A. J., Grecis, R. K., Jeltsch, A., Hajkova, P. and Sarkies, P. (2018) 'Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity', *Nature Genetics*. Nature Publishing Group, 50(3), pp. 452–459.
- Rossant, J. and Tam, P. P. L. (2009) 'Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse.', *Development (Cambridge, England)*. The Company of Biologists Ltd, 136(5), pp. 701–13.
- Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, A., Plachot, M., Paldi, A. and Viegas-Pequignot, E. (1998) 'Chromosome methylation patterns during mammalian preimplantation development', *Genes & Development*. Cold Spring Harbor Laboratory Press, 12(14), pp. 2108–2113.

Roundtree, I. A., Evans, M. E., Pan, T. and He, C. (2017) 'Dynamic RNA Modifications in Gene Expression Regulation', *Cell*, 169(7), pp. 1187–1200.

Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A. J. R., Carlton, M. B. L., Pearce, J. J., Barton, S. C., Surani, M. A., Ryan, K., Nehls, M. C., Wilson, V. and Evans, M. J. (2000) 'Eomesodermin is required for mouse trophoblast development and mesoderm formation', *Nature*. Nature Publishing Group, 404(6773), pp. 95–99.

Saijo, Y., Sato, G., Usui, K., Sato, M., Sagawa, M., Kondo, T., Minami, Y. and Nukiwa, T. (2001) 'Expression of nucleolar protein p120 predicts poor prognosis in patients with stage I lung adenocarcinoma', *Ann. Onc.*, 12(8), pp. 1121–1125.

Saitou, M., Kagiwada, S. and Kurimoto, K. (2011) 'Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells', *Development*, 139(1).

Sakita-Suto, S., Kanda, A., Suzuki, F., Sato, S., Takata, T. and Tatsuka, M. (2007) 'Aurora-B regulates RNA methyltransferase NSUN2.', *Molecular biology of the cell*, 18(3), pp. 1107–17.

Santoemma, G. (2018) 'Recent methodologies for studying the soil organic matter', *Applied Soil Ecology*. Elsevier, 123, pp. 546–550.

Santos, F., Peat, J., Burgess, H., Rada, C., Reik, W. and Dean, W. (2013) 'Active demethylation in mouse zygotes involves cytosine deamination and base excision repair.', *Epigenetics & chromatin*, 6(1), p. 39.

Santos, J., Filipe Pereira, C., Di-Gregorio, A., Spruce, T., Alder, O., Rodriguez, T., Azuara, V., Merckenschlager, M. and Fisher, A. G. (2010) 'Differences in the epigenetic and reprogramming properties of pluripotent and extra-embryonic stem cells implicate chromatin remodelling as an important early event in the developing mouse embryo', *Epigenetics & Chromatin*. BioMed Central, 3(1), p. 1.

Sasaki, K., Yokobayashi, S., Nakamura, T., Okamoto, I., Yabuta, Y., Kurimoto, K., Ohta, H., Moritoki, Y., Iwatani, C., Tsuchiya, H., Nakamura, S., Sekiguchi, K., Sakuma, T., Yamamoto, Takashi, Mori, T., Wolffen, K., Nakagawa, M., Yamamoto, Takuya, Takahashi, K., et al. (2015) 'Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells.', *Cell stem cell*. Elsevier, 17(2), pp. 178–94.

Saunders, C. L., Chiodini, B. D., Sham, P., Lewis, C. M., Abkevich, V., Adeyemo, A. A., de Andrade, M., Arya, R., Berenson, G. S., Blangero, J., Boehnke, M., Borecki, I. B., Chagnon, Y. C., Chen, W., Comuzzie, A. G., Deng, H.-W., Duggirala, R., Feitosa, M. F., Froguel, P., et al. (2007) 'Meta-Analysis of Genome-wide Linkage Studies in BMI and Obesity\*', *Obesity*. John Wiley & Sons, Ltd, 15(9), pp. 2263–2275.

Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M. and Lyko, F. (2010) 'RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage.', *Genes & development*, 24(15), pp. 1590–5.

Schier, A. F. (2007) 'The maternal-zygotic transition: death and birth of RNAs.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 316(5823), pp. 406-7.

Schmidt, C. S., Bultmann, S., Meilinger, D., Zacher, B., Tresch, A., Maier, K. C., Peter, C., Martin, D. E., Leonhardt, H. and Spada, F. (2012) 'Global DNA Hypomethylation Prevents Consolidation of Differentiation Programs and Allows Reversion to the Embryonic Stem Cell State', *PLoS ONE*. Edited by J. G. Knott. Public Library of Science, 7(12), p. e52629.

Schossere, M., Minois, N., Angerer, T. B., Amring, M., Dellago, H., Harreither, E., Calle-Perez, A., Pircher, A., Gerstl, M. P., Pfeifenberger, S., Brandl, C., Sonntagbauer, M., Kriegner, A., Linder, A., Weinhäusel, A., Mohr, T., Steiger, M., Mattanovich, D., Rinnerthaler, M., Karl, T., Sharma, S. Entian, K. D., Kos, M., Breitenbach, M., Wilson, I. B.H., Polacek, N., Grillari-Voglauer, R., Breitenbach-Koller, L. and Grillari, J. (2015) 'Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan', *Nature Communications*. Nature Publishing Group, 6, p. 6158.

Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., Sanjana, N. E., Freinkman, E., Pacold, M. E., Satija, R., Mikkelsen, T. S., Hacohen, N., Zhang, F., Carr, S. A., Lander, E. S., *et al.* (2014) 'Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites.', *Cell reports*, 8(1), pp. 284-96.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W. and Reik, W. (2012) 'The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells', *Molecular Cell*. Cell Press, 48(6), pp. 849-862.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M. and Matsui, Y. (2005) 'Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice', *Developmental Biology*. Academic Press, 278(2), pp. 440-458.

Sharma, S., Yang, J., Watzinger, P., Kötter, P. and Entian, K.-D. (2013) 'Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively.', *Nucleic acids research*. Oxford University Press, 41(19), pp. 9062-76.

Sheaffer, K. L., Kim, R., Aoki, R., Elliott, E. N., Schug, J., Burger, L., Schübeler, D. and Kaestner, K. H. (2014) 'DNA methylation is required for the control of stem cell differentiation in the small intestine.', *Genes & development*, 28(6), pp. 652-64.

Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., Kumagai, S., Ochiai, K., Suzuki, T. and Igarashi, K. (2017) 'S-Adenosylmethionine Synthesis Is Regulated by Selective N6-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1', *Cell Reports*. Cell Press, 21(12), pp. 3354-3363.

Sheth, U. and Parker, R. (2003) 'Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies', *Science*. American Association for the Advancement of Science, 300(5620), pp. 805–808.

Skourti-Stathaki, K., Kamieniarz-Gdula, K. and Proudfoot, N. J. (2014) 'R-loops induce repressive chromatin marks over mammalian gene terminators', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 516(7531), pp. 436–439.

Smallwood, S. A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S. R. and Kelsey, G. (2011) 'Dynamic CpG island methylation landscape in oocytes and preimplantation embryos', *Nature Genetics*. Nature Publishing Group, 43(8), pp. 811–814.

Smemo, S., Tena, J. J., Kim, K.-H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., Aneas, I., Credidio, F. L., Sobreira, D. R., Wasserman, N. F., Lee, J. H., Puviindran, V., Tam, D., Shen, M., Son, J. E., Vakili, N. A., Sung, H.-K., Naranjo, S., Acemel, R. D., *et al.* (2014) 'Obesity-associated variants within FTO form long-range functional connections with IRX3', *Nature*. Nature Publishing Group, 507(7492), pp. 371–375.

Smith, A. (2017) 'Formative pluripotency: the executive phase in a developmental continuum.', *Development (Cambridge, England)*. Oxford University Press for The Company of Biologists Limited, 144(3), pp. 365–373.

Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A. and Meissner, A. (2012) 'A unique regulatory phase of DNA methylation in the early mammalian embryo', *Nature*. Nature Publishing Group, 484(7394), pp. 339–344.

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., Suter, C. M. and Preiss, T. (2012) 'Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA.', *Nucleic acids research*, 40(11), pp. 5023–33.

Stadtfeld, M. and Hochedlinger, K. (2010) 'Induced pluripotency: history, mechanisms, and applications.', *Genes & development*. Cold Spring Harbor Laboratory Press, 24(20), pp. 2239–63.

Stratigopoulos, G., Padilla, S. L., LeDuc, C. A., Watson, E., Hattersley, A. T., McCarthy, M. I., Zeltser, L. M., Chung, W. K. and Leibel, R. L. (2008) 'Regulation of *Fto/Ftm* gene expression in mice and humans', *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*. American Physiological Society, 294(4), pp. R1185–R1196.

Strobel, M. C. and Abelson, J. (1986) 'Effect of intron mutations on processing and function of *Saccharomyces cerevisiae* SUP53 tRNA in vitro and in vivo.', *Molecular and cellular biology*. American Society for Microbiology, 6(7), pp. 2663–73.

Su, D., Chan, C. T. Y., Gu, C., Lim, K. S., Chionh, Y. H., McBee, M. E., Russell, B. S., Babu, I. R., Begley, T. J. and Dedon, P. C. (2014) 'Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry.', *Nature protocols*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 9(4), pp. 828–41.

Sun, Z., Terragni, J., Borgaro, J. G., Liu, Y., Yu, L., Guan, S., Wang, H., Sun, D., Cheng, X., Zhu, Z., Pradhan, S. and Zheng, Y. (2013) 'High-Resolution Enzymatic Mapping of Genomic 5-Hydroxymethylcytosine in Mouse Embryonic Stem Cells', *Cell Reports*. Cell Press, 3(2), pp. 567–576.

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

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. and Rao, A. (2009) 'Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.', *Science (New York, N.Y.)*, 324(5929), pp. 930–5.

Tam, P. P., Tan, S. S., Takagi, N. and Heasman, J. (1992) 'The somitogenetic potential of cells in the primitive streak and the tail bud of the organogenesis-stage mouse embryo.', *Development (Cambridge, England)*. The Company of Biologists Ltd, 115(3), pp. 703–15.

Tsuda, H., Maynard-Currie, C. E., Reid, L. H., Yoshida, T., Edamura, K., Maeda, N., Smithies, O. and Jakobovits, A. (1997) 'Inactivation of the Mouse HPRT Locus by a 203-bp Retroposon Insertion and a 55-kb Gene-Targeted Deletion: Establishment of New HPRT-Deficient Mouse Embryonic Stem Cell Lines', *Genomics*, 42(3), pp. 413–421.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H. R., Nakayama, J. and Okano, M. (2006) 'Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b', *Genes to Cells*. Blackwell Publishing Inc, 11(7), pp. 805–814.

Tung, Y.-C. L., Ayuso, E., Shan, X., Bosch, F., O'Rahilly, S., Coll, A. P. and Yeo, G. S. H. (2010) 'Hypothalamic-Specific Manipulation of Fto, the Ortholog of the Human Obesity Gene FTO, Affects Food Intake in Rats', *PLoS ONE*. Edited by A. Bartolomucci. Public Library of Science, 5(1), p. e8771.

Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G. and Lyko, F. (2012) 'RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis.', *Nature structural & molecular biology*. Nature Publishing Group, 19(9), pp. 900–5.

Turner, D. A., Trott, J., Hayward, P., Rué, P. and Martinez Arias, A. (2014) 'An interplay between extracellular signalling and the dynamics of the exit from pluripotency drives cell fate decisions in mouse ES cells.', *Biology open*. The Company of Biologists Ltd, 3(7), pp. 614–26.

- Uchida, S., Hara, K., Kobayashi, A., Funato, H., Hobara, T., Otsuki, K., Yamagata, H., McEwen, B. S. and Watanabe, Y. (2010) 'Early Life Stress Enhances Behavioral Vulnerability to Stress through the Activation of REST4-Mediated Gene Transcription in the Medial Prefrontal Cortex of Rodents', *Journal of Neuroscience*, 30(45), pp. 15007–15018.
- van Deemter, J. J., Zuiderweg, F. J. and Klinkenberg, A. (1956) 'Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography', *Chemical Engineering Science*. Pergamon, 5(6), pp. 271–289.
- Vazquez-Zapien, G. J., Mata-Miranda, M. M., Sanchez-Monroy, V., Delgado-Macuill, R. J., Perez-Ishiwara, D. G. and Rojas-Lopez, M. (2016) 'FTIR Spectroscopic and Molecular Analysis during Differentiation of Pluripotent Stem Cells to Pancreatic Cells.', *Stem cells international*. Hindawi Limited, 2016, p. 6709714.
- Vincent, S. D., Dunn, N. R., Sciammas, R., Shapiro-Shalef, M., Davis, M. M., Calame, K., Bikoff, E. K. and Robertson, E. J. (2005) 'The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse', *Development*. The Company of Biologists Ltd, 132(6), pp. 1315–1325.
- WADDINGTON, C. H. (1957) 'The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser.', *The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser*. London: George Allen & Unwin, Ltd. Available at: <https://www.cabdirect.org/cabdirect/abstract/19580101706> (Accessed: 21 October 2018).
- C.H. Waddington, C.H. (1942) 'The epigenotype.' *Endeavour*, 1 (1942), pp. 18-20
- Wallrath, L. L. and Elgin, S. C. (1995) 'Position effect variegation in Drosophila is associated with an altered chromatin structure.', *Genes & development*. Cold Spring Harbor Laboratory Press, 9(10), pp. 1263–77.
- Wang, F., Minakhina, S., Tran, H., Changela, N., Kramer, J. and Steward, R. (2018) 'Tet protein function during Drosophila development', *PLOS ONE*. Edited by S. Pradhan. Public Library of Science, 13(1), p. e0190367.
- Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., Zou, T. and Yin, P. (2016) 'Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex', *Nature*. Nature Publishing Group, 534(7608), pp. 575–578.
- 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. and He, C. (2014) 'N6-methyladenosine-dependent regulation of messenger RNA stability', *Nature*. Nature Publishing Group, 505(7481), pp. 117–120.

- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H. and He, C. (2015) 'N6-methyladenosine Modulates Messenger RNA Translation Efficiency', *Cell*, 161(6), pp. 1388–1399.
- Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z. and Zhao, J. C. (2014) 'N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells.', *Nature cell biology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 16(2), pp. 191–8.
- WATSON, J. D. and CRICK, F. H. C. (1953) 'Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid', *Nature*. Nature Publishing Group, 171(4356), pp. 737–738.
- Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., Shi, H., Cui, X., Su, R., Klungland, A., Jia, G., Chen, J. and He, C. (2018) 'Differential m6A, m6Am, and m1A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm', *Molecular Cell*. Cell Press, 71(6), pp. 973-985.e5.
- Wen, J., Zeng, Y., Fang, Z., Gu, J., Ge, L., Tang, F., Qu, Z., Hu, J., Cui, Y., Zhang, K., Wang, J., Li, S., Sun, Y. and Jin, Y. (2017) 'Single-cell analysis reveals lineage segregation in early post-implantation mouse embryos.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 292(23), pp. 9840–9854.
- Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L. and Owen-Hughes, T. (1999) 'Nucleosome mobilization catalysed by the yeast SWI/SNF complex', *Nature*. Nature Publishing Group, 400(6746), pp. 784–787.
- Williams, K., Christensen, J., Pedersen, M. T., Johansen, J. V., Cloos, P. A. C., Rappsilber, J. and Helin, K. (2011) 'TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity', *Nature*. Nature Publishing Group, 473(7347), pp. 343–348.
- Wu, Q., Niebuhr, E., Yang, H. and Hansen, L. (2005) 'Determination of the “critical region” for cat-like cry of Cri-du-chat syndrome and analysis of candidate genes by quantitative PCR.', *European journal of human genetics : EJHG*, 13(4), pp. 475–85.
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., Sun, H.-Y., Li, A., Ping, X.-L., Lai, W.-Y., Wang, X., Ma, H.-L., Huang, C.-M., Yang, Y., Huang, N., Jiang, G.-B., Wang, H.-L., Zhou, Q., Wang, X.-J., Zhao, Y.-L., Yang, Y.-G. (2016) 'Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing', *Molecular Cell*. Cell Press, 61(4), pp. 507–519.
- Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E. and Carrington, J. C. (2004) 'Genetic and Functional Diversification of Small RNA Pathways in Plants', *PLoS Biology*. Edited by Detlef Weigel. Public Library of Science, 2(5), p. e104.

- Xing, J., Yi, J., Cai, X., Tang, H., Liu, Z., Zhang, X., Martindale, J. L., Yang, X., Jiang, B., Gorospe, M. and Wang, W. (2015) 'NSun2 Promotes Cell Growth via Elevating Cyclin-Dependent Kinase 1 Translation.', *Molecular and cellular biology*, 35(23), pp. 4043-52.
- Yagi, M., Kishigami, S., Tanaka, A., Semi, K., Mizutani, E., Wakayama, S., Wakayama, T., Yamamoto, T. and Yamada, Y. (2017) 'Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation', *Nature. Nature Research*, 548(7666), p. 224. doi: 10.1038/nature23286.
- Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y. and Saitou, M. (2008) 'Critical function of Prdm14 for the establishment of the germ cell lineage in mice', *Nature Genetics*. Nature Publishing Group, 40(8), pp. 1016-1022.
- Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T. and Handa, H. (2006) 'P-TEFb-Mediated Phosphorylation of hSpt5 C-Terminal Repeats Is Critical for Processive Transcription Elongation', *Molecular Cell*. Cell Press, 21(2), pp. 227-237.
- Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y. and Saitou, M. (2008) 'Critical function of Prdm14 for the establishment of the germ cell lineage in mice', *Nature Genetics*. Nature Publishing Group, 40(8), pp. 1016-1022.
- Yamasaki, S., Ivanov, P., Hu, G. and Anderson, P. (2009) 'Angiogenin cleaves tRNA and promotes stress-induced translational repression', *The Journal of Cell Biology*. Rockefeller University Press, 185(1), pp. 35-42.
- Yan, M., Wang, Y., Hu, Y., Feng, Y., Dai, C., Wu, J., Wu, D., Zhang, F. and Zhai, Q. (2013) 'A high-throughput quantitative approach reveals more small RNA modifications in mouse liver and their correlation with diabetes.', *Analytical chemistry*. American Chemical Society, 85(24), pp. 12173-81.
- Yang, X., Yang, Y., Sun, B.-F., Chen, Y.-S., Xu, J.-W., Lai, W.-Y., Li, A., Wang, X., Bhattarai, D. P., Xiao, W., Sun, H.-Y., Zhu, Q., Ma, H.-L., Adhikari, S., Sun, M., Hao, Y.-J., Zhang, B., Huang, C.-M., Huang, N., Jiang, G.-Bin., Zhao, Y.-L., Wang, H.-L., Sun, Y.-P. and Yang, Y.-G. (2017) '5-methylcytosine promotes mRNA export – NSUN2 as the methyltransferase and ALYREF as an m5C reader', *Cell Research*. Nature Publishing Group, 27(5), pp. 606-625.
- Ying, Q.-L., Wray, J., Nichols, J., Battle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008) 'The ground state of embryonic stem cell self-renewal.', *Nature*. Nature Publishing Group, 453(7194), pp. 519-23.
- Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y. I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Schöler, H. R. and Matsui, Y. (1999) 'Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice.', *Development, growth & differentiation*, 41(6), pp. 675-84.

- Yost, R. A. and Enke, C. G. (1979) 'Triple Quadrupole Mass Spectrometry', *Analytical Chemistry*. American Chemical Society, 51(12), pp. 1251A-1264A.
- Young, R. A. (2011) 'Control of the Embryonic Stem Cell State', *Cell*. Cell Press, 144(6), pp. 940-954.
- Zauri, M., Berridge, G., Thézénas, M.-L., Pugh, K. M., Goldin, R., Kessler, B. M. and Kriaucionis, S. (2015) 'CDA directs metabolism of epigenetic nucleosides revealing a therapeutic window in cancer', *Nature*. Nature Publishing Group, 524(7563), pp. 114-118.
- Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., Chen, Y., Sulman, E. P., Xie, K., Böglér, O., Majumder, S., He, C. and Huang, S. (2017) 'm6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program', *Cancer Cell*. Cell Press, 31(4), pp. 591-606.e6.
- Zhang, T., Chen, P., Li, W., Sha, S., Wang, Y., Yuan, Z., Shen, B. and Chen, L. (2019) 'Cognitive deficits in mice lacking Nsun5, a cytosine-5 RNA methyltransferase, with impairment of oligodendrocyte precursor cells', *Glia*, 67(4), pp. 688-702.
- Zhang, X., Liu, Z., Yi, J., Tang, H., Xing, J., Yu, M., Tong, T., Shang, Y., Gorospe, M. and Wang, W. (2012) 'The tRNA methyltransferase NSun2 stabilizes p16INK4 mRNA by methylating the 3'-untranslated region of p16', *Nature Communications*. Nature Publishing Group, 3(1), p. 712.
- Zhang, Y. and Reinberg, D. (2001) 'Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.', *Genes & development*. Cold Spring Harbor Laboratory Press, 15(18), pp. 2343-60.
- Zhang, Y. and Sachs, M. S. (2015) 'Control of mRNA Stability in Fungi by NMD, EJC and CBC Factors Through 3'UTR Introns.', *Genetics*, 200(4), pp. 1133-48.
- 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., *et al.* (2013) 'ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility', *Molecular Cell*. Cell Press, 49(1), pp. 18-29.
- Zhou, Q., Li, T. and Price, D. H. (2012) 'RNA Polymerase II Elongation Control', *Annual Review of Biochemistry*. Annual Reviews , 81(1), pp. 119-143.
- Zhu, H., Shyh-Chang, N., Segrè, A. V., Shinoda, G., Shah, S. P., Einhorn, W. S., Takeuchi, A., Engreitz, J. M., Hagan, J. P., Kharas, M. G., Urbach, A., Thornton, J. E., Triboulet, R., Gregory, R. I., Altshuler, D. and Daley, G. Q. (2011) 'The Lin28/let-7 Axis Regulates Glucose Metabolism', *Cell*. Cell Press, 147(1), pp. 81-94.

## Appendices

| Page No. | Type of work: | Name of work                     | Source of work                                                                                                                                                                                                                                                   | Copyright holder and contact                                                                                     | permission requested on | I have permission yes /no | Permission note                                                                                                                                         |
|----------|---------------|----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-------------------------|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
|          | Figure        | Fig. 1.1 The evolution of RNMTs. | Jurkowski, T. P. and Jeltsch, A. (2011) 'On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2.', <i>PloS one</i> . Public Library of Science, 6(11), p. e28104.                                                                             | PLOS ONE<br><a href="mailto:onepress@plos.org">onepress@plos.org</a>                                             | N/A                     | Yes                       | Creative Commons Attribution (CC-BY) Licence<br><a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a> |
|          | Figure        | Fig. 1.1 The evolution of RNMTs. | Bujnicki, J. M., Feder, M., Ayres, C. L. and Redman, K. L. (2004) 'Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases.', <i>Nucleic acids research</i> , 32(8), pp. 2453-63. | © 2019 Oxford University Press<br><a href="mailto:journals.permissions@oup.com">journals.permissions@oup.com</a> | 31/07/2019              | Yes                       | Licence provided below                                                                                                                                  |

|  |        |                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                    |            |     |                                                                                                                                                         |
|--|--------|---------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | Figure | Fig. 1.3 Morphology of Trdmt1 and Nsun2 double knockout.            | Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G. and Lyko, F. (2012) 'RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis.', <i>Nature structural &amp; molecular biology</i> . Nature Publishing Group, 19(9), pp. 900–5.                                                                           | © 2019 Springer Nature Publishing AG<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a> | 31/07/2019 | Yes | Licence provided below                                                                                                                                  |
|  | Figure | Fig. 1.4 The reported roles of RNMTs in transcriptional regulation. | Cheng, J. X., Chen, L., Li, Y., Cloe, A., Yue, M., Wei, J., Watanabe, K. A., Shammo, J. M., Anastasi, J., Shen, Q. J., Larson, R. A., He, C., Le Beau, M. M. and Vardiman, J. W. (2018) 'RNA cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia', <i>Nature Communications</i> . Nature Publishing Group, 9(1), p. 1163. | © 2019 Springer Nature Publishing AG<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a> | N/A        | Yes | Creative Commons Attribution (CC-BY) Licence<br><a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a> |

|  |        |                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                 |            |     |                                                                                                                                                               |
|--|--------|-----------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|------------|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | Figure | Fig 1.6<br>Reversible m6A<br>methylation.                                                           | Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., Dai, Q., Chen, W. and He, C. (2014) 'A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation.', <i>Nature chemical biology</i> . Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 10(2), pp. 93–5. | © 2019 Springer<br>Nature Publishing<br>AG<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a>        | 05/08/2019 | Yes | Licence provided below                                                                                                                                        |
|  | Figure | Fig 1.7 RNA<br>modification<br>groups<br>transcripts for<br>cellular<br>processes.                  | Roundtree, I. A., Evans, M. E., Pan, T. and He, C. (2017) 'Dynamic RNA Modifications in Gene Expression Regulation', <i>Cell</i> , 169(7), pp. 1187–1200.                                                                                                                                                                                                        | Copyright © 2019<br>Elsevier<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a>                      | 05/08/2019 | Yes | Licence provided below                                                                                                                                        |
|  | Figure | Fig 1.8<br>YTHDF2 is<br>required to clear<br>maternal<br>mRNAs and<br>permit zygotic<br>development | Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P. N., Monahan, J., Carrieri, C., Enright, A. J. and O'Carroll, D. (2017) 'The RNA m6A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the                                                                                                                           | © 2017 The<br>Authors.<br>Published by<br>Elsevier Inc.<br><a href="mailto:donal.ocarroll@ed.ac.uk">donal.ocarroll@ed.ac.uk</a> | N/A        | Yes | Creative Commons<br>Attribution (CC-BY)<br>Licence<br><a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a> |

|  |        |                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                              |            |     |                                              |
|--|--------|--------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|------------|-----|----------------------------------------------|
|  |        | beyond the 2-cell stage.                               | Maternal Transcriptome and Oocyte Competence', <i>Molecular Cell</i> . Cell Press, 67(6), pp. 1059-1067.e4.                                                                                                                                                                                                                                                                                                         |                                                                                                              |            |     |                                              |
|  | Figure | Fig. 1.9 The mRNA of core stem cell genes contain m6A. | 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., <i>et al.</i> (2014) 'm(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells.', <i>Cell stem cell</i> . Elsevier, 15(6), pp. 707-19. | Copyright © 2014 Elsevier Inc.<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a> | 05/08/2019 | Yes | Licence provided below                       |
|  | Figure | Fig. 1.9 The mRNA of core stem cell genes contain m6A. | Young, R. A. (2011) 'Control of the Embryonic Stem Cell State', <i>Cell</i> . Cell Press, 144(6), pp. 940-954.                                                                                                                                                                                                                                                                                                      | Copyright © 2011 Elsevier Inc.<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a> | 05/08/2019 | Yes | Licence provided below                       |
|  | Figure | Fig 1.10 Developmental progression of pluripotency in  | Smith, A. (2017) 'Formative pluripotency: the executive phase in a developmental continuum.', <i>Development</i> (Cambridge,                                                                                                                                                                                                                                                                                        | © 2019 The Company of Biologists Ltd                                                                         | N/A        | Yes | Creative Commons Attribution (CC-BY) Licence |

|  |        |                                                                            |                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                 |            |     |                                                                                                         |
|--|--------|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|------------|-----|---------------------------------------------------------------------------------------------------------|
|  |        | mouse and human embryos.                                                   | <i>England</i> ). Oxford University Press for The Company of Biologists Limited, 144(3), pp. 365–373.                                                                                                                                                                                                                                                                | <a href="mailto:permissions@biologists.com">permissions@biologists.com</a>                                      |            |     | <a href="https://creativecommons.org/licenses/by/3.0/">https://creativecommons.org/licenses/by/3.0/</a> |
|  | Figure | Fig 1.11 Waddington's Classical Epigenetic Landscape.                      | Goldberg, A. D., Allis, C. D. and Bernstein, E. (2007) 'Epigenetics: A Landscape Takes Shape', <i>Cell</i> . Cell Press, 128(4), pp. 635–638.                                                                                                                                                                                                                        | Copyright © 2007 Elsevier Inc. <a href="mailto:customercare@copyright.com">customercare@copyright.com</a>       | 05/08/2019 | Yes | Licence provided below                                                                                  |
|  | Figure | Fig. 1.12 A schematic of mouse pre-implantation and germ cell development. | Saitou, M., Kagiwada, S. and Kurimoto, K. (2011) 'Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells', <i>Development</i> , 139(1). doi: 10.1242/dev.142679                                                                                                                                                                    | © 2019 The Company of Biologists Ltd <a href="mailto:permissions@biologists.com">permissions@biologists.com</a> | 05/08/2019 | Yes | Licence provided below                                                                                  |
|  | Figure | Fig 3.1 Eddy diffusion in a packed chromatography column.                  | Harvy, D. (2014) Contributions to Band Broadening in Chromatography. Available at: <a href="https://community.asdlib.org/imageandvideoexchangeforum/2013/08/01/contributions-to-band-broadening-in-chromatography">https://community.asdlib.org/imageandvideoexchangeforum/2013/08/01/contributions-to-band-broadening-in-chromatography</a> [Accessed 2 jul. 2019]. | Image and video exchange forum                                                                                  | N/A        | Yes | Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0)                    |

|  |        |                                                                                          |                                                                                                                                                            |                                                                                                       |                |     |                        |
|--|--------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|----------------|-----|------------------------|
|  |        |                                                                                          |                                                                                                                                                            |                                                                                                       |                |     |                        |
|  | Figure | Fig 3.2<br>Quadrupole<br>mass<br>spectrometer.<br>Diagram of<br>instrument<br>operation. | Santoiemma, G. (2018) 'Recent<br>methodologies for studying the<br>soil organic matter', <i>Applied Soil<br/>Ecology</i> . Elsevier, 123, pp. 546-<br>550. | © 2017 Elsevier<br>B.V.<br><a href="mailto:customercare@copyright.com">customercare@copyright.com</a> | 05/08/201<br>9 | Yes | Licence provided below |

**Appendix table 1** Permission notes for cited figures.

## OXFORD UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS

Jul 31, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number 4639441414405

License date Jul 31, 2019

Licensed content publisher Oxford University Press

Licensed content publication Nucleic Acids Research

Licensed content title Sequence-structure-function studies of tRNA:m5C

methyltransferase Trm4p and its relationship to DNA:m5C and

RNA:m5U methyltransferases

Licensed content author Bujnicki, Janusz M.; Feder, Marcin

Licensed content date Apr 15, 2004

Type of Use Thesis/Dissertation

Institution name

Title of your work RNA Methylation in Epigenetic Reprogramming

Publisher of your work Imperial College London

Expected publication date Oct 2019

Permissions cost 0.00 USD

Value added tax 0.00 USD

Total 0.00 USD

Title RNA Methylation in Epigenetic Reprogramming

Institution name Imperial College London

Expected presentation date Oct 2019

Portions Figure 6

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Publisher Tax ID GB125506730

Total 0.00 USD

Terms and Conditions

### STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details.
2. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12.
3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it  
7/31/2019 RightsLink Printable License  
<https://s100.copyright.com/AppDispatchServlet> 2/2  
apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.
4. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint.
5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.
6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from [www.oxfordjournals.org](http://www.oxfordjournals.org) Should there be a problem clearing these rights, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)
7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.

8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

12. Other Terms and Conditions:

v1.4

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

## SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Jul 31, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number 4639440478301

License date Jul 31, 2019

Licensed Content Publisher Springer Nature

Licensed Content Publication Nature Structural & Molecular Biology

Licensed Content Title RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis

Licensed Content Author Francesca Tuorto, Reinhard Liebers, Tanja Musch, Matthias Schaefer, Sarah Hofmann et al.

Licensed Content Date Aug 12, 2012

Licensed Content Volume 19

Licensed Content Issue 9

Type of Use Thesis/Dissertation

Requestor type academic/university or research institute

Format print and electronic

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

High-res required no

Will you be translating? no

Circulation/distribution <501

Author of this Springer

Nature content

no

Title RNA Methylation in Epigenetic Reprogramming

Institution name Imperial College London

Expected presentation date Oct 2019

Portions Figure 1b and 1e

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Total 0.00 GBP

Terms and Conditions

### Springer Nature Customer Service Centre GmbH

#### Terms and Conditions

7/31/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 2/4

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

#### 1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

#### 2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

**2. 3.** Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to [Journalpermissions@springernature.com/bookpermissions@springernature.com](mailto:Journalpermissions@springernature.com/bookpermissions@springernature.com) for these rights.

**2. 4.** Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

**2. 5.** An alternative scope of licence may apply to signatories of the [STM Permissions Guidelines](#), as amended from time to time.

### **3. Duration of Licence**

**3. 1.** A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

#### **Scope of**

#### **Licence Duration of Licence**

Post on a website 12 months

Presentations 12 months

Books and

journals

Lifetime of the edition in the language

purchased

### **4. Acknowledgement**

**4. 1.** The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the

7/31/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 3/4

figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

### **5. Restrictions on use**

**5. 1.** Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

**5. 2.** You must not use any Licensed Material as part of any design or trademark.

**5. 3.** Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

### **6. Ownership of Rights**

**6. 1.** Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

### **7. Warranty**

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND

WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

### **8. Limitations**

**8. 1. BOOKS ONLY:**Where 'reuse in a dissertation/thesis' has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline ([www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/)).

### **9. Termination and Cancellation**

**9. 1.** Licences will expire after the period shown in Clause 3 (above).

**9. 2.** Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

## Appendix 1 — Acknowledgements:

7/31/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 4/4

### For Journal Content:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### For Advance Online Publication papers:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

### For Adaptations/Translations:

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### For Advance Online Publication papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

### For Book content:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc)] [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)]

### Other Conditions:

Version 1.2

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

## SPRINGER NATURE LICENSE TERMS AND CONDITIONS

Aug 05, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number 4642471177054

License date Aug 05, 2019

Licensed Content Publisher Springer Nature

Licensed Content Publication Nature Chemical Biology

Licensed Content Title A METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation

Licensed Content Author Jianzhao Liu, Yanan Yue, Dali Han, Xiao Wang, Ye Fu et al.

Licensed Content Date Dec 6, 2013

Licensed Content Volume 10

Licensed Content Issue 2

Type of Use Thesis/Dissertation

Requestor type academic/university or research institute

Format print and electronic

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

High-res required no

Will you be translating? no

Circulation/distribution <501

Author of this Springer

Nature content

no

Title RNA Methylation in Epigenetic Reprogramming

Institution name Imperial College London

Expected presentation date Oct 2019

Portions Figure 1

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Total 0.00 GBP

Terms and Conditions

### Springer Nature Customer Service Centre GmbH

#### Terms and Conditions

8/5/2019 RightsLink Printable License

https://s100.copyright.com/AppDispatchServlet 2/4

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

#### 1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

#### 2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly

excluded from the licence.

**2. 3.** Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to [Journalpermissions@springernature.com/bookpermissions@springernature.com](mailto:Journalpermissions@springernature.com/bookpermissions@springernature.com) for these rights.

**2. 4.** Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

**2. 5.** An alternative scope of licence may apply to signatories of the [STM Permissions Guidelines](#), as amended from time to time.

### **3. Duration of Licence**

**3. 1.** A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

#### **Scope of**

#### **Licence Duration of Licence**

Post on a website 12 months

Presentations 12 months

Books and

journals

Lifetime of the edition in the language

purchased

### **4. Acknowledgement**

**4. 1.** The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 3/4

figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

### **5. Restrictions on use**

**5. 1.** Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

**5. 2.** You must not use any Licensed Material as part of any design or trademark.

**5. 3.** Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

### **6. Ownership of Rights**

**6. 1.** Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

### **7. Warranty**

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND

WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

### **8. Limitations**

**8. 1. BOOKS ONLY:**Where 'reuse in a dissertation/thesis' has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline ([www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/)).

### **9. Termination and Cancellation**

**9. 1.** Licences will expire after the period shown in Clause 3 (above).

**9. 2.** Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

## Appendix 1 — Acknowledgements:

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 4/4

### For Journal Content:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### For Advance Online Publication papers:

Reprinted by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

### For Adaptations/Translations:

Adapted/Translated by permission from [the Licensor]: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [the Licensor]: on behalf of Cancer Research UK: : [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication)]

### For Advance Online Publication papers:

Reprinted by permission from The [the Licensor]: on behalf of Cancer Research UK: [Journal Publisher (e.g. Nature/Springer/Palgrave)] [JOURNAL NAME] [REFERENCE CITATION (Article name, Author(s) Name), [COPYRIGHT] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)]

### For Book content:

Reprinted/adapted by permission from [the Licensor]: [Book Publisher (e.g. Palgrave Macmillan, Springer etc)] [Book Title] by [Book author(s)] [COPYRIGHT] (year of publication)

### Other Conditions:

Version 1.2

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

## ELSEVIER LICENSE TERMS AND CONDITIONS

Aug 05, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4642460428931

License date Aug 05, 2019

Licensed Content Publisher Elsevier

Licensed Content Publication Cell

Licensed Content Title Dynamic RNA Modifications in Gene Expression Regulation

Licensed Content Author Ian A. Roundtree, Molly E. Evans, Tao Pan, Chuan He

Licensed Content Date Jun 15, 2017

Licensed Content Volume 169

Licensed Content Issue 7

Licensed Content Pages 14

Start Page 1187

End Page 1200

Type of Use reuse in a thesis/dissertation

Intended publisher of new

work

other

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

Format both print and electronic

Are you the author of this

Elsevier article?

No

Will you be translating? No

Original figure numbers Figure 3

Title of your

thesis/dissertation

RNA Methylation in Epigenetic Reprogramming

Publisher of new work Imperial College London

Expected completion date Oct 2019

Estimated size (number of

pages)

1

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Publisher Tax ID GB 494 6272 12

Total 0.00 GBP

Terms and Conditions

## INTRODUCTION

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 2/5

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

## GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of

chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 3/5

terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied

permissions.

## LIMITED LICENSE

The following terms and conditions apply only to specific license types:

**15. Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

**16. Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at

<http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

**17. For journal authors:** the following clauses are applicable in addition to the above:

### Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peerreviewed,

nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes authorincorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:  
immediately

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 4/5

via their non-commercial person homepage or blog

by updating a preprint in arXiv or RePEc with the accepted manuscript

via their research institute or institutional repository for internal institutional

uses or as part of an invitation-only research collaboration work-group

directly by providing copies to their students or to research collaborators for their personal use

for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

After the embargo period

via non-commercial hosting platforms such as their institutional repository

via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

link to the formal publication via its DOI

bear a CC-BY-NC-ND license - this is easy to do

if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

### **Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 5/5

party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

### **Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

### **Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the

Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

Associating advertising with the full text of the Article

Charging fees for document delivery or access

Article aggregation

Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

## **20. Other Conditions:**

v1.9

Questions? [customer care@copyright.com](mailto:customer care@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

## ELSEVIER LICENSE TERMS AND CONDITIONS

Aug 05, 2019

---

This Agreement between Imperial College London -- Karl Hudson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

|                                              |                                                                                                                                                                                                                                                      |
|----------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| License Number                               | 4642451504248                                                                                                                                                                                                                                        |
| License date                                 | Aug 05, 2019                                                                                                                                                                                                                                         |
| Licensed Content Publisher                   | Elsevier                                                                                                                                                                                                                                             |
| Licensed Content Publication                 | Cell Stem Cell                                                                                                                                                                                                                                       |
| Licensed Content Title                       | m6A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells                                                                                                                                                                 |
| Licensed Content Author                      | Pedro J. Batista,Benoit Molinie,Jinkai Wang,Kun Qu,Jiajing Zhang,Lingjie Li,Donna M. Bouley,Ernesto Lujan,Bahareh Haddad,Kaveh Daneshvar,Ava C. Carter,Ryan A. Flynn,Chan Zhou,Kok-Seong Lim,Peter Dedon,Marius Wernig,Alan C. Mullen,Yi Xing et al. |
| Licensed Content Date                        | Dec 4, 2014                                                                                                                                                                                                                                          |
| Licensed Content Volume                      | 15                                                                                                                                                                                                                                                   |
| Licensed Content Issue                       | 6                                                                                                                                                                                                                                                    |
| Licensed Content Pages                       | 13                                                                                                                                                                                                                                                   |
| Start Page                                   | 707                                                                                                                                                                                                                                                  |
| End Page                                     | 719                                                                                                                                                                                                                                                  |
| Type of Use                                  | reuse in a thesis/dissertation                                                                                                                                                                                                                       |
| Portion                                      | figures/tables/illustrations                                                                                                                                                                                                                         |
| Number of figures/tables/illustrations       | 1                                                                                                                                                                                                                                                    |
| Format                                       | both print and electronic                                                                                                                                                                                                                            |
| Are you the author of this Elsevier article? | No                                                                                                                                                                                                                                                   |
| Will you be translating?                     | No                                                                                                                                                                                                                                                   |
| Original figure numbers                      | Figure 1b                                                                                                                                                                                                                                            |
| Title of your thesis/dissertation            | RNA Methylation in Epigenetic Reprogramming                                                                                                                                                                                                          |
| Publisher of new work                        | Imperial College London                                                                                                                                                                                                                              |
| Expected completion date                     | Oct 2019                                                                                                                                                                                                                                             |

|                                  |                                                                                                                                        |
|----------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Estimated size (number of pages) | 1                                                                                                                                      |
| Requestor Location               | Imperial College London<br>MRC LMS, Hammersmith Hospital<br><br>London, W12 0NN<br>United Kingdom<br><br>Attn: Imperial College London |
| Publisher Tax ID                 | GB 494 6272 12                                                                                                                         |
| Total                            | 0.00 GBP                                                                                                                               |
| Terms and Conditions             |                                                                                                                                        |

## INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

## GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically

authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

13. **Revocation:** Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

#### LIMITED LICENSE

The following terms and conditions apply only to specific license types:

14. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

15. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

16. **For journal authors:** the following clauses are applicable in addition to the above:

#### Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-

incorporated changes suggested during submission, peer review and editor-author communications

Authors can share their accepted author manuscript:

- ♦ immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- ♦ After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- ♦ link to the formal publication via its DOI
- ♦ bear a CC-BY-NC-ND license - this is easy to do
- ♦ if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

17. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you

scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

18. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

**Terms & Conditions applicable to all Open Access articles published with Elsevier:** Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

**Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- ♦ Associating advertising with the full text of the Article
- ♦ Charging fees for document delivery or access
- ♦ Article aggregation
- ♦ Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

Other Conditions:

v1.9

**Questions?** [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

---

**ELSEVIER LICENSE  
TERMS AND CONDITIONS**

Aug 05, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4642461123132

License date Aug 05, 2019

Licensed Content Publisher Elsevier

Licensed Content Publication Cell

Licensed Content Title Control of the Embryonic Stem Cell State

Licensed Content Author Richard A. Young

Licensed Content Date Mar 18, 2011

Licensed Content Volume 144

Licensed Content Issue 6

Licensed Content Pages 15

Start Page 940

End Page 954

Type of Use reuse in a thesis/dissertation

Intended publisher of new

work

other

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

Format both print and electronic

Are you the author of this

Elsevier article?

No

Will you be translating? No

Original figure numbers Figure 6

Title of your

thesis/dissertation

RNA Methylation in Epigenetic Reprogramming

Publisher of new work Imperial College London

Expected completion date Oct 2019

Estimated size (number of

pages)

1

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Publisher Tax ID GB 494 6272 12

Total 0.00 GBP

Terms and Conditions

**INTRODUCTION**

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet 2/5>

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

**GENERAL TERMS**

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of

chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet/3/5>

terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied

permissions.

## LIMITED LICENSE

The following terms and conditions apply only to specific license types:

**15. Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

**16. Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at

<http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

**17. For journal authors:** the following clauses are applicable in addition to the above:

### Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peerreviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes authorincorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

immediately

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet/4/5>

via their non-commercial person homepage or blog

by updating a preprint in arXiv or RePEc with the accepted manuscript

via their research institute or institutional repository for internal institutional

uses or as part of an invitation-only research collaboration work-group

directly by providing copies to their students or to research collaborators for their personal use

for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

After the embargo period

via non-commercial hosting platforms such as their institutional repository

via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

link to the formal publication via its DOI

bear a CC-BY-NC-ND license - this is easy to do

if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

### **Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet/5/5>

party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

### **Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

### **Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the

user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

Associating advertising with the full text of the Article

Charging fees for document delivery or access

Article aggregation

Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

## 20. Other Conditions:

v1.9

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

## **ELSEVIER LICENSE TERMS AND CONDITIONS**

Aug 05, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4642470240234

License date Aug 05, 2019

Licensed Content Publisher Elsevier

Licensed Content Publication Applied Soil Ecology

Licensed Content Title Recent methodologies for studying the soil organic matter

Licensed Content Author Giacomo Santoemma

Licensed Content Date Feb 1, 2018

Licensed Content Volume 123

Licensed Content Issue n/a

Licensed Content Pages 5

Start Page 546

End Page 550

Type of Use reuse in a thesis/dissertation

Intended publisher of new

work

other

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

Format both print and electronic

Are you the author of this

Elsevier article?

No

Will you be translating? No

Original figure numbers Figure 3

Title of your

thesis/dissertation

RNA Methylation in Epigenetic Reprogramming

Publisher of new work Imperial College London

Expected completion date Oct 2019

Estimated size (number of  
pages)

1

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Publisher Tax ID GB 494 6272 12

Total 0.00 GBP

Terms and Conditions

## **INTRODUCTION**

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

## **GENERAL TERMS**

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment

terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. **Revocation:** Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

### **LIMITED LICENSE**

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

#### **Preprints:**

A preprint is an author's own write-up of research results and analysis, it has not been peerreviewed,

nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available

version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:  
immediately

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 4/5

via their non-commercial person homepage or blog

by updating a preprint in arXiv or RePEc with the accepted manuscript

via their research institute or institutional repository for internal institutional

uses or as part of an invitation-only research collaboration work-group

directly by providing copies to their students or to research collaborators for their personal use

for private scholarly sharing as part of an invitation-only work group on

commercial sites with which Elsevier has an agreement

After the embargo period

via non-commercial hosting platforms such as their institutional repository

via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

link to the formal publication via its DOI

bear a CC-BY-NC-ND license - this is easy to do

if aggregated with other manuscripts, for example in a repository or other site, be

shared in alignment with our hosting policy not be added to or enhanced in any way to

appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not

allowed to download and post the published electronic version of your chapter, nor may you

scan the printed edition to create an electronic version. **Posting to a repository:** Authors are

permitted to post a summary of their chapter only in their institution's repository.

**19. Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

### **Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet> 5/5

party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

### **Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

### **Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

Associating advertising with the full text of the Article

Charging fees for document delivery or access

Article aggregation

Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

**20. Other Conditions:**

v1.9

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or  
+1-978-646-2777.

**ELSEVIER LICENSE  
TERMS AND CONDITIONS**

Aug 05, 2019

This Agreement between Imperial College London -- Karl Hudson ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4642470586839

License date Aug 05, 2019

Licensed Content Publisher Elsevier

Licensed Content Publication Cell

Licensed Content Title Epigenetics: A Landscape Takes Shape

Licensed Content Author Aaron D. Goldberg,C. David Allis,Emily Bernstein

Licensed Content Date Feb 23, 2007

Licensed Content Volume 128

Licensed Content Issue 4

Licensed Content Pages 4

Start Page 635

End Page 638

Type of Use reuse in a thesis/dissertation

Intended publisher of new

work

other

Portion figures/tables/illustrations

Number of

figures/tables/illustrations

1

Format both print and electronic

Are you the author of this

Elsevier article?

No

Will you be translating? No

Original figure numbers Figure 1

Title of your

thesis/dissertation

RNA Methylation in Epigenetic Reprogramming

Publisher of new work Imperial College London

Expected completion date Oct 2019

Estimated size (number of  
pages)

1

Requestor Location Imperial College London

MRC LMS, Hammersmith Hospital

London, W12 0NN

United Kingdom

Attn: Imperial College London

Publisher Tax ID GB 494 6272 12

Total 0.00 GBP

Terms and Conditions

**INTRODUCTION**

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

**GENERAL TERMS**

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at [permissions@elsevier.com](mailto:permissions@elsevier.com)). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment

terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. **Revocation:** Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

### **LIMITED LICENSE**

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com>; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at <http://www.elsevier.com>. All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve:** In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

#### **Preprints:**

A preprint is an author's own write-up of research results and analysis, it has not been peerreviewed,

nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available

version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:  
immediately

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet 4/5>

via their non-commercial person homepage or blog

by updating a preprint in arXiv or RePEc with the accepted manuscript

via their research institute or institutional repository for internal institutional

uses or as part of an invitation-only research collaboration work-group

directly by providing copies to their students or to research collaborators for their personal use

for private scholarly sharing as part of an invitation-only work group on

commercial sites with which Elsevier has an agreement

After the embargo period

via non-commercial hosting platforms such as their institutional repository

via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

link to the formal publication via its DOI

bear a CC-BY-NC-ND license - this is easy to do

if aggregated with other manuscripts, for example in a repository or other site, be

shared in alignment with our hosting policy not be added to or enhanced in any way to

appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a [CrossMark logo](#), the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's [posting policy](#) for further information.

18. **For book authors** the following clauses are applicable in addition to the above:

Authors are permitted to place a brief summary of their work online only. You are not

allowed to download and post the published electronic version of your chapter, nor may you

scan the printed edition to create an electronic version. **Posting to a repository:** Authors are

permitted to post a summary of their chapter only in their institution's repository.

**19. Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

### **Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third

8/5/2019 RightsLink Printable License

<https://s100.copyright.com/AppDispatchServlet 5/5>

party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our [open access license policy](#) for more information.

### **Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

### **Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by/4.0>.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at <http://creativecommons.org/licenses/by-nc-nd/4.0>.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

Associating advertising with the full text of the Article

Charging fees for document delivery or access

Article aggregation

Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

**20. Other Conditions:**

v1.9

Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

1

**Hudson, Karl**

**From:** permissions <permissions@biologists.com>

**Sent:** 06 August 2019 08:02

**To:** Hudson, Karl; permissions

**Subject:** RE: Permission to use an extract for a PhD Thesis

Dear Karl,

Thank you for your request. Permission is granted for a thesis without charge. We cannot however grant

as a CC BY-NC licence as the use of our images/figures is tightly controlled.

The acknowledgement is very important and should state "reproduced/adapted with permission" and give

the source journal name - the acknowledgement should either provide full citation details or refer to the

relevant citation in the article reference list - the full citation details should include authors, journal,

year, volume, issue and page citation.

Where appearing online or in other electronic media, a link should be provided to the original article

(e.g. via DOI).

Development: <http://www.biologists.com/development>

We wish you the best of luck with your thesis.

Kind regards

Richard

**Richard Grove**

Commercial Manager

The Company of Biologists Ltd

Bidder Building, Station Road, Histon, Cambridge, CB24 9LF, UK

T: +44 (0) 1223 632 850 | [richard.grove@biologists.com](mailto:richard.grove@biologists.com) | [www.biologists.com](http://www.biologists.com)

Registered office: The Company Of Biologists Ltd, Bidder Building, Station Road, Histon, Cambridge CB24 9LF,

United Kingdom, Registered in England and Wales. Company Limited by Guarantee No 514735.

Registered Charity

No 277992 The information contained in this message and any attachment is confidential, legally privileged and is

intended for the addressee only. Any dissemination, distribution, copying, disclosure or use of this message/attachment or its contents is strictly prohibited and may be unlawful. No contract is intended or implied,

unless confirmed by hard copy. If you have received this message in error, please inform the sender and delete it

from your mailbox or any other storage mechanism. The Company of Biologists Ltd cannot accept liability for any

statements made which are clearly the senders' own and not expressly made on behalf of The Company of

Biologists Ltd or one of their agents.

**From:** Hudson, Karl <k.hudson15@lms.mrc.ac.uk>

**Sent:** 05 August 2019 15:35

**To:** permissions <permissions@biologists.com>

**Subject:** Permission to use an extract for a PhD Thesis

05/08/2019

2

Dear The Company of Biologists

I am completing my PhD thesis at Imperial College London entitled 'RNA Methylation in Epigenetic

Reprogramming'.

I seek your permission to reprint, in my thesis an extract from: Saitou, M., Kagiwada, S. and Kurimoto, K.

(2011) 'Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells',

*Development*, 139(1). The extract to be reproduced is: Figure 1.

I would like to include the extract in my thesis which will be added to Spiral, Imperial's institutional repository

<http://spiral.imperial.ac.uk/> and made available to the public under a Creative Commons Attribution-Non

Commercial 4.0 International Licence (CC BY-NC)

If you are happy to grant me all the permissions requested, please return a signed copy of this letter. If you

wish to grant only some of the permissions requested, please list these and then sign.

Yours sincerely,

Name: Mr Karl Hudson, MBiochem

Signed: KHudson

Mail: [karl.hudson15@imperial.ac.uk](mailto:karl.hudson15@imperial.ac.uk)

Permission granted for the use requested above:

I confirm that I am the copyright holder of the extract above and hereby give permission to include it in your

thesis which will be made available, via the internet, for non-commercial purposes under the terms of the user

licence.

[please edit the text above if you wish to grant more specific permission]

Signed:

Name:

Organisation:

Job title: