Single-Cell N^6-Methyladenosine (m6A) Mapping in Oocytes and Preimplantation Embryos

Yanjiao Li

Thesis for the Degree of Philosophiae Doctor (PhD)

Department of Microbiology
Oslo University Hospital
and
Department of Molecular Medicine
Institute of Basic Medical Sciences
University of Oslo

Oslo, Norway
2021

UNIVERSITY OF OSLO

Oslo University Hospital

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-348-0048-1

Cover: Hanne Baadsgaard Utigard.
Print production: Graphics Center, University of Oslo.

Table of Contents

Acknowledgments

The work presented in this thesis was conducted at the Department of Microbiology, Oslo University Hospital (OUS), Rikshospitalet. The PhD candidate is also affiliated with the Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo (UiO).

Undertaking the PhD has been a truly wonderful experience for me, and it would not have been possible to complete without valuable support and guidance from many people.

First, I would like to express my sincere gratitude to my main supervisor, Arne Klungland, for his continuous support during my PhD study and for his patience, inspiration, and immense knowledge. Second, I would like to express my deep gratitude to my cosupervisor John Arne Dahl for all the guidance, support, and encouragement that he gave me. He helped me through my research and writing of the thesis. His expertise was invaluable in formulating the research questions and methodology. Furthermore, his scientific way of thinking pushed me to sharpen my thinking, and this will benefit me throughout my life. I cannot imagine better supervisors than you.

I greatly appreciate the support that I received from Yunhao Wang and Kin Fai Au. Thank you very much for providing excellent support with bioinformatics analysis. I am also grateful to all those people who contributed to my PhD work, especially to Maria, Linda, Erik, and Leif. Many thanks to Trine, Gareth, and Peter at the Department of Reproductive Medicine, OUS, for their assistance throughout my study. Furthermore, I am grateful to Ingunn and Shaista at Norwegian Transgenic Center (NTS) for their contributions to sample collection and technical support. My deep appreciation goes out to my local colleagues as well: Madeleine, Sherif, Adeel, Mari, Stig Ove and Magnar. Thank you very much for your guidance in instrumentation and techniques. I also want to give a huge thank you to all the lab members and administrative staff for creating a supportive atmosphere during my PhD work. Many thanks to my past and current office members too for stimulating discussions during my study. I gratefully acknowledge my PhD funding received from UiO:Life Science convergence environments.

Last but not least, I would like to thank my family members for supporting me spiritually throughout my work with this thesis. There are no words to convey how much I love you. The only way is that I will work harder to repay your trust and love that I received.

Yanjiao Li

Oslo, 2021

List of papers

Paper I.

Li Y [#], Wang Y [#], Rodriguez MV, Lindeman LC, Skuggen LE, Rasmussen E, Jermstad I, Khan S, Fosslie ML, Skuland T, Klemsdal EK, Dalen KT, Fedorscak P, Greggains GD, **Klungland A**[*], Au KF[*], **Dahl JA**[*]. Single-cell MeRIP-seq maps m6A in oocytes and embryos. *Under revision to* ***Nature Biotechnology***

Paper II.

Li Y, Wang Y, Rodriguez MV, Skuggen LE, Au KF, **Klungland A**[*]**, Dahl JA**[*]. Transcriptome-wide mapping of N^6-methyladenosine via antibody-based immunoprecipitation and high-throughput sequencing at ultra-low/single cell level. *Under submission*

Paper III.

Wang Y[#], **Li Y**[#], Skuland T, Zhou CJ, Jermstad I, Khan S, Li A, Hashim A, Dalen KT, Greggains GD, **Klungland A**[*]**, Dahl JA**[*], Au KF[*]. The landscape of m6A in mammalian preimplantation embryos (Alternative title: The RNA m6A landscape in mouse oocytes and preimplantation embryos). *Under submission*

Paper IV.

Zhang Z, Manaf A, **Li Y**, Perez SP, Suganthan R, Dahl JA, Bjørås M & **Klungland A**[*]. Histone Methylations Define Neural Stem/Progenitor Cell Subtypes in the Mouse *Molecular Neurobiology*, volume 57, pages 997–1008 (2020)

# Equal contributions

* Corresponding author

List of abbreviations

carRNAs	Chromosome-associated regulatory RNAs
m6A	N^6-methyladenosine
MZT	Maternal to zygotic transition
HTS	High-throughput sequencing
single-cell MeRIP-seq/m6A-seq	scMeRIP-seq
TSSs	Transcription start sites
m6Am	2'-O-dimethyladenosine
m1A	N^1-methyladenosine
RRBS-Seq	Reduced-representation bisulfite sequencing
m5C	5-methylcytosine
hm5C	5-hydroxymethylcytosine
hm5C	5-hydroxymethylcytosine
A-to-I editing	Adenosine-to-inosine editing
dsRNA	Double-stranded RNA
ADARs	Adenosine deaminases acting on RNA
C-to-U editing	Cytosine-to-uracil editing
Ψ	Pseudouridine
m7G	N^7-methylguanosine
METTL 3	Methyltransferase like 3
WTAP	Wilms Tumor 1 Associated Protein
Xist	Long non-coding RNA X-inactive specific transcript
mTORC1	Mechanistic target of rapamycin complex 1
FTO	Fat mass and obesity-associated protein
ALKBH5	AlkB homolog 5
YTH	YT521-B homology
GV	Germinal vesicle
ESCs	Embryonic stem cells
ERVs	Endogenous retroviruses
ZGA	Zygotic genome activation
IPed	Immunoprecipitated
4SU	4-thiouridine
RT	Reverse transcription
SAM	S-adenosine methionine
a6A	N^6-allyladenosine
dm6A	N^6-dithiolsitolmethyladenosine
SeAdoYn	SAM cofactor surrogate
TGS	Third generation sequencing, or long-read sequencing
PacBio	Pacific Biosciences
SMRT sequencing	Single Molecule Real Time sequencing
ONT sequencing	Oxford Nanopore Technologies sequencing
MaLRs	Mammalian apparent LTRs retrotransposons
MII oocyte	Metaphase II oocyte
IF staining	Immunofluorescence staining
LAF bench	Laminar flow bench
rRNAs	Ribosomal RNAs
MMLV-RT	Moloney murine leukemia virus reverse transcriptase
TEs	Transposable elements or repeat RNAs
LINEs	Long interspersed nuclear elements
SINEs	Short interspersed nuclear elements
LTRs	Long terminal repeat elements

Summary

In English:

N^6-methyladenosine (m6A) is the most prevalent and abundant internal modification in mRNAs. It plays regulatory roles that control gene expression in diverse cellular processes, including RNA transcription, splicing, nuclear export, decay, and translation. Transcriptome-level m6A mapping has become an essential tool in innumerable m6A studies. Despite the considerable improvements in m6A mapping methodology, the applicability of current methods is limited by the need for large amounts of input materials. In this thesis, we first developed an m6A ultra-low/single-cell MeRIP-seq (m(6)A-specific methylated RNA immunoprecipitation with next-generation sequencing, it is also abbreviated as scMeRIP-seq) mapping assay to study transcriptome-wide m6A modification profiles using limited numbers of cells and single cells. We describe the whole protocol in a step-by-step manner. In brief, m6A-containing RNA is immunoprecipitated with an anti-m6A antibody prior to single-tube library construction and sequencing with a high-throughput DNA sequencer, constituting an easy, robust and reliable technique for mapping m6A at the ultra-low/single-cell level. Second, we validated the method by profiling m6A modifications in serial dilutions of poly(A)-selected RNA from mouse liver and demonstrated its utility by assays single zebrafish zygotes, single mouse MII oocytes and mouse blastocysts. Intriguingly, we found that m6A was deposited predominantly on the MTA_Mm retrotransposon in MII oocytes. MTA_Mm is a unique retrotransposon element for reprogramming of the genome during the maternal-to-zygotic transition. Third, by applying scMeRIP–seq in mouse oocytes and preimplantation embryos in different stages of development, we investigated the landscapes of m6A in early mammalian development, revealing unique spatiotemporal m6A dynamics that accompany early mammalian development. Collectively, our work opens a new avenue for studying m6A in a single cell and in rare cells in a transcriptome-wide manner.

På norsk:

N6-metyladenosin (m6A) er den mest utbredte og anrikede interne mRNA-modifikasjonen. Den har regulatoriske roller som kontrollerer genuttrykk i forskjellige cellulære prosesser, inkludert RNA-transkripsjon, spleising, transport ut av cellekjernen, nedbryting og translasjon. m6A-kartlegging på transkriptomnivå har blitt et viktig verktøy i utallige m6A-studier. Til tross for betydelige forbedringer i m6A-kartleggingsmetodikk, er anvendeligheten av dagens metoder begrenset av behovet for store mengder cellemateriale. I denne avhandlingen utviklet vi først en m6A ultra-lav/enkeltcelle MeRIP-seq (m(6)A-spesifikk metylert RNA-immunutfelling etterfulgt av neste generasjonssekvensering, som forkortet heter scMeRIP-seq) kartleggingsanalyse for å studere m6A-modifikasjonsprofiler på transkripsjonsbredt nivå ved bruk av et begrenset antall celler og enkeltceller. Vi beskriver hele protokollen steg for steg. Kort fortalt blir RNA som har m6A-modifikasjonen immunutfelt med et anti-m6A-antistoff før bibliotek klargjøres i ett og samme rør, og deretter følger sekvensering ved hjelp av en DNA-sekvenseringsmaskin med høy kapasitet. Dette er en enkel, robust og pålitelig teknikk for å kartlegge m6A på ultra-lavt/enkeltcellenivå. Videre validerte vi metoden ved å profilere m6A-modifikasjoner i seriefortynninger av poly(A)-selektert RNA fra muselever og demonstrerte dens nytte ved analyser av enkeltzygoter fra sebrafisk, og enkelt-oocytter og blastocyster fra mus. Interessant nok fant vi at m6A hovedsakelig blir avsatt på MTA_Mm retrotransposoner i MII-oocytter. MTA_Mm er et unikt retrotransposonelement for omprogrammering av genomet under overgangen hvor utviklingen går fra å være under maternell kontroll til å kontrolleres av zygotens egne genprodukter. Til slutt undersøkte vi m6A-landskapet i tidlig pattedyrutvikling ved å bruke scMeRIP-seq på museoocytter og preimplantasjonsembryoer i forskjellige utviklingsstadier, og med dette avslørte vi en unik romlig-temporal m6A-dynamikk som finner sted under tidlig pattedyrutvikling. Samlet sett åpner arbeidet vårt opp for en ny måte å studere m6A på i enkeltceller og i sjeldne celler på et transkripsjonsbredt nivå.

Introduction

The life of a mammalian embryo begins with a sperm fertilizing an oocyte. After fertilization, the zygote acquires the capability to become a new organism. It divides and differentiates in an organized manner. The central dogma of molecular biology describes the flow of genetic information from DNA to RNA and from RNA to protein via processes called transcription and translation, respectively (Crick 1970). The epigenome is the sum of the chemical modifications on DNA and histones as well as certain chromosome-associated regulatory RNAs (carRNAs) that affect the transcriptional activity of the genome without altering the DNA sequence (Klemm et al. 2019). Modifications on DNA and histone residues regulate transcription, and discoveries of the dynamic regulation of such marks have led to tremendous progress in our understanding of gene regulation. Analogous to DNA methylation in epigenetics, RNAs are subject to over 100 distinct posttranscriptional modifications that regulate various processes, including RNA stability and decay and protein translation, and these mechanisms are collectively termed epitranscriptome (Helm and Motorin 2017). Compared to the discovery of dynamic DNA and histone modifications, the discovery of reversible RNA modifications is much more recent, and the understanding of the epitranscriptome is still in its infancy. However, over the past decade, novel methods for mapping RNA modifications have been widely used and have become an effective tool for exploring the epitranscriptome (Zaccara et al. 2019).

During gametogenesis and embryogenesis, epigenomic and epitranscriptomic information is coded and transmitted to the next generation. However, due to the scarcity of embryos, it is unrealistic to use a large quantity of materials to explore the preimplantation development mechanism. In the past decade, the rapid advancement of sequencing technologies has enabled detailed profiling of the genomic and epigenomic code during early embryogenesis. RNA sequencing is the most widely used method. Novel methodologies allow detailed studies on DNA methylation (Guo et al. 2014; Zhu et al. 2018), chromatin accessibility (Gao et al. 2018; Wu et al. 2018), histone modifications (Dahl et al. 2016; Zhang et al. 2016), the chromosome conformation (Collombet et al. 2020), etc. N6-methyladenosine (m6A) is the most common internal modification of mRNA and affects multiple biological processes related to mRNA metabolism (Roundtree et al. 2017a; Zhao et al. 2017a; Zaccara et al. 2019). Dynamic m6A modification has key roles in many processes occurring during reproduction, including gametogenesis, maternal-to-zygotic transition (MZT) and early embryo development (Lasman et al. 2020b; Zhang et al. 2020). To be noted, m6A is also required for morphogenesis in plants and is key to post-embryonic development (Zhong et al. 2008; Arribas-Hernandez et al. 2018;

Arribas-Hernandez et al. 2020). Regarding m6A and its function in model organisms and cultured stem cells have been completed in recent years, yet the precise roles of m6A in gametogenesis and embryogenesis are not well understood.

In 2012, two groups described a novel strategy for sequence-specific and transcriptome-wide m6A profiling, MeRIP-seq/m6A-seq (Dominissini et al. 2012; Meyer et al. 2012; Zaccara et al. 2019). MeRIP-seq/m6A-seq exploited an m6A-specific antibody to immunoprecipitate m6A-containing RNAs and construct a library compatible with high-throughput sequencing (HTS). Later, an expanding panel of m6A mapping methods were developed. To date, more than 2000 m6A methylomes have been deposited in databases, providing valuable information on how m6A influences diverse physiological processes and diseases. Despite the power of these m6A mapping methods, many challenges remain to be addressed, partly due to methodological challenges. The large amount of input RNA required for immunoprecipitation and library preparation, making it impossible for application in early embryos and cancer biopsies. This thesis describes the development of ultra-low/single-cell MeRIP-seq (scMeRIP-seq) and its application in oocytes and early embryos.

RNA modifications

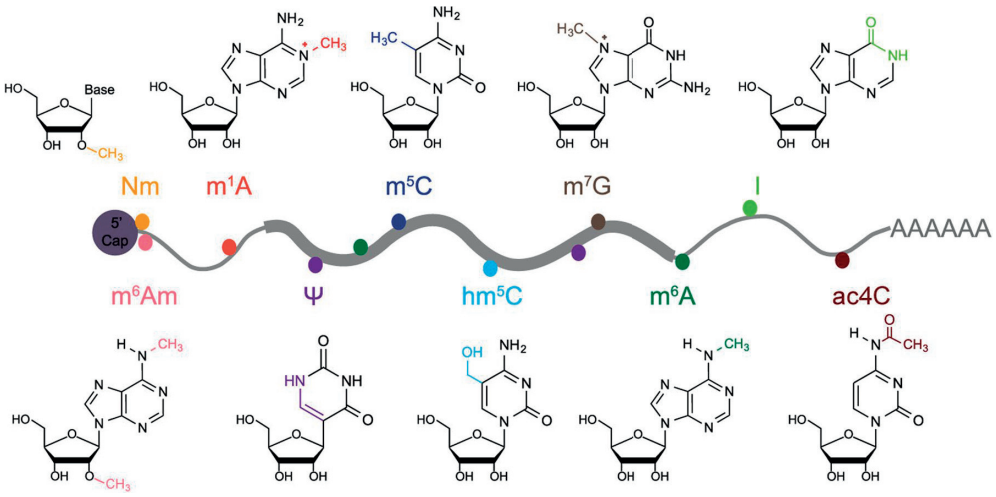


Figure 1. Known mRNA Modifications. m^6A, N^6-methyladenosine; m^1A, N^1-methyladenosine; m^6Am, N^6,2′-O-dimethyladenosineinosine; m^5C, 5-methylcytosine; hm^5C, 5-hydroxymethylcytosine; Nm, 2′-O-methylated nucleotides; Ψ, pseudouridine; I, inosine; m^7G, N^7-methylguanosine; ac4C, N^4-acetylcytidine. The figure is from (Song and Yi 2019).

The basic characteristics of RNA have been known for more than half a century; it is a linear molecule composed of four ribonucleotide bases: guanine, uracil, adenine, and cytosine (denoted by the letters G, U, A, and C, respectively). Compared with studies on DNA and

protein modifications, studies on RNA modifications have been underappreciated (Wiener and Schwartz 2021). During the past decade, advances in technology have facilitated the discovery of diverse RNA modifications, and more than 160 types of RNA modifications have been identified and recorded in the database (**Figure 1**) (Boccaletto et al. 2018). Modifications make RNA more rigid and more flexible without changing the ribonucleotides; these modifications influence RNA structure, processing, stability, protein interactions, etc. (Boccaletto et al. 2018). These diverse RNA modifications have been assigned to a new layer of gene regulation, which is termed 'epitranscriptomics' or the 'epitranscriptome' (Saletore et al. 2012; Frye et al. 2016).

Methyladenosine Modifications

m6A, N^1-methyladenosine (m1A), 2'-O-dimethyladenosine (m6Am)

m6A is the most prevalent internal modification in eukaryotic mRNA. The m6A/A ratio in polyadenylated RNA has been calculated to be approximately 0.4% by LC–MS/MS (Liu et al. 2014). The METTL3–METTL14 writer complex mediates N6-adenosine methylation of nuclear RNA (Liu et al. 2014), and recent studies have identified several subunits in addition to methyltransferases (Zaccara et al. 2019; Zhao et al. 2020). The FTO and ALKBH5 (termed 'erasers') demethylases convert m6A to A, and the discovery of their activity was a key discovery that demonstrated the reversibility of RNA modifications in eukaryotic cells (Jia et al. 2011; Zheng et al. 2013a). In 2012, two m6A mapping methods were developed to map m6A across the transcriptome and have been used in innumerable studies.

The first two studies utilized antibodies to enrich m6A-containing RNA and analyzed this RNA by massively parallel sequencing. However, antibody-dependent strategies recognize not only m6A but also 2'-O-dimethyladenosine (m6Am) (Linder et al. 2015). The m6Am modification is mainly localized at transcription start sites (TSSs), the first nucleotide proximal to the 5′ cap, and the m6Am/A ratio is approximately 0.02–0.05% (Wei et al. 2018; Sun et al. 2019). m6Am is a dynamic and reversible epitranscriptomic modification that determines mRNA stability (Mauer et al. 2017).

Similar to the findings for m6A, transcriptome-wide identification of N^1-methyladenosine (m1A) showed its widespread presence at thousands of sites (Dominissini et al. 2016; Xiao et al. 2016). However, the amount of m1A is much lower than that of m6A, and the m1A/A ratio in polyadenylated RNA is approximately 0.02%. Later, the anti-m1A antibody was used to combine with misincorporation-based reverse transcription and detected only a handful of m1A sites (Li et al. 2017b; Safra et al. 2017). TRMT10C and TRMT61B were identified as

mitochondria-localizing m1A methyltransferases (Li et al. 2017b; Safra et al. 2017). The discrepancy between studies is due to a combination of SNPs, alignment, sequencing errors, and computational considerations that led to sites being misinterpreted as m1A sites in earlier studies (Schwartz 2018; Sas-Chen and Schwartz 2019). An optimal evolved reverse transcriptase enabled the detection of well-characterized m1A sites and revealed hundreds of m1A sites in human mRNA (Zhou et al. 2019a). In conclusion, these data suggest that m1A could be present in a small number of sites in mRNA.

Cytosine Modifications

5-methylcytosine (m5C), 5-hydroxymethylcytosine (hm5C), N^4-acetylcytidine (ac4C)

5-Methylcytosine (m5C) is present on 0.03–0.1% of cytosines in RNA (Huber et al. 2015; Legrand et al. 2017), and studies have identified ~10,000 m5C sites in human mRNA (Schaefer et al. 2009; Squires et al. 2012; Amort et al. 2017; Yang et al. 2017). In addition, mRNAs were found to be very sparsely modified by m5C or not methylated at all in mouse embryonic stem cells (Legrand et al. 2017). several hundred exonic m5C sites were identified in human and mouse tissues; approximately 62–70% of the sites had low methylation levels (<20% methylation), while 8–10% of the sites were moderately or highly methylated (>40% methylation) (Huang et al. 2019b). Analysis of HeLa cell RNA by reduced representation bisulfite sequencing (RRBS-Seq) revealed hundreds of m5C sites in noncoding RNAs and mRNAs (Khoddami et al. 2019; Schumann et al. 2020). Collectively, these findings support the presence of m5C on mammalian mRNA, although the stoichiometry is context-dependent. NSUN2 is a methyltransferase that catalyzes the conversion of cytosine to 5-methylcytosine (m5C) at a CRGGG (where R stands for A or G) motif (Huang et al. 2019b; Schumann et al. 2020).

The 5-hydroxymethylcytosine (hm5C) level is much lower than the m5C level, found on approximately 0.001–0.003% of cytosines (Huber et al. 2015). An antibody-based method identified over 3000 hm5C peaks in the Drosophila brain, preferentially at polyadenylated RNAs and located at the CDS region (Delatte et al. 2016). High levels of hm5C modification in the brainstem, cerebellum, and hippocampus in mice have been identified by dot blot analysis (Miao et al. 2016).

Inosine

Adenosine-to-inosine (A-to-I) editing refers to enzymatic deamination of adenosine nucleosides and results in their permanent change to inosine in RNA (Bass 2002; Nishikura 2010). It is a co/posttranscriptional modification and mainly occurs in double-stranded RNA (dsRNA) (Bass and Weintraub 1988). This modification was identified based on sequence comparison between the DNA sequences and mRNA sequences of a particular glutamate receptor (Sommer et al. 1991). RNA editing can be catalyzed by dsRNA-specific adenosine deaminases acting on RNA (ADARs) (Polson et al. 1991). Three members of this family are encoded by the mammalian genome: ADAR1, ADAR2 (also called ADARB1) and ADAR3 (also called ADARB2) (Eisenberg and Levanon 2018). A-to-I modifications account for nearly 90% of all editing events in RNA, and more than one million editing sites have been reported (Eisenberg and Levanon 2018). The development of high-throughput sequencing has allowed the identification of various edited bases, such as those that have undergone the rare cytosine-to-uracil (C-to-U) editing. However, at present, data on noncanonical sites might be contaminated by mismapped reads or sequencing errors (Piskol et al. 2013)

Pseudouridine (Ψ)

Ψ is possibly the most abundant base modification except for m6A, and the ratio of Ψ/U in mRNA is approximately 0.2–0.6% (Li et al. 2015). Three groups developed similar chemical-based profiling methods for Ψ, termed pseudo-seq, Ψ-seq or PSI-seq, that depend on blocking of reverse transcriptase at Ψ modifications and permit studies at single-base resolution (Carlile et al. 2014; Lovejoy et al. 2014; Schwartz et al. 2014a). These methods do not contain a step for pre-enrichment of Ψ-containing RNAs; thus, enrichment of low-abundance Ψ modifications is lost. More recently, a method with selective chemical labeling and pre-enrichment was developed for profiling pseudouridylated RNAs (Li et al. 2015). Another quantitative method for profiling Ψ sites at true base-pair resolution revealed several hundred Ψ sites (Khoddami et al. 2019). Ψ sites depend on conserved cognate pseudouridine synthases that are conserved from yeast to mammals (Schwartz et al. 2014a). Collectively, these studies support the existence of hundreds to thousands of Ψ sites in mRNA. However, further methodologies for Ψ profiling with quantification and higher sensitivity are still needed, and the sequencing depth, bioinformatics algorithms and the cell-dependent context need to be considered for better confirmation of the presence of Ψ sites on mRNA.

Guanosine methylation

Guanosines in RNA can also be methylated. N^7-methylguanosine (m7G) is installed at the 5′ cap cotranscriptionally during transcription initiation (Cowling 2009). This m7G cap modification stabilizes transcripts to protect against exonucleolytic degradation and modulates transcription elongation, pre-mRNA splicing, nuclear export, and translation (Ramanathan et al. 2016; Song et al. 2020). Recently, some internal m7G modifications were identified in mRNAs and miRNAs (Pandolfini et al. 2019; Zhang et al. 2019a), and the internal m7G/G level was found to be approximately 0.002%–0.05% in mammalian mRNAs (Chu et al. 2018; Zhang et al. 2019a). It should be noted that detection of mRNA from *Escherichia coli* or yeast did not identify any internal m7G modifications (Enroth et al. 2019).

m6A modification

History of m6A discovery

The existence of m6A-methylated messenger RNA was detected in two studies in 1974 (Desrosiers et al. 1974; Perry and Kelley 1974). Shortly thereafter, a study reported that cytoplasmic simian virus exhibited a relatively large amount of m6A (Lavi and Shatkin 1975). In 1975, two independent studies also showed that m6A was present in HeLa cell mRNA (Furuichi et al. 1975) and in mouse myeloma cell RNA (Adams and Cory 1975). In 1978, m6A was shown to be associated with mRNA instability in HeLa cells (Sommer et al. 1978b). In 1997, m6A was found to be catalyzed by a complex multicomponent enzyme and the subunit (methyltransferase 3, METTL3) of human mRNA m6A methyltransferase was identified (Bokar et al. 1997). Later, two groups separately revealed that METTL3 homologs were also present in yeast and Arabidopsis and that their mutation resulted in specific developmental arrest (Clancy et al. 2002; Zhong et al. 2008). However, it is not possible to rule out the possibility that this m6A came from contamination—i.e., that the detected m6A came from contaminating RNA species such as tRNA, which is also highly modified. In 2012, two independent studies developed methods of antibody-based immunoprecipitation followed by high-throughput sequencing to map the m6A RNA methylome at the whole-transcriptome level (Dominissini et al. 2012; Meyer et al. 2012). It was found that more than 10,000 m6A peaks exist in mouse and human mRNAs.

The abundance and functions of m6A on RNA are determined by multiple proteins, namely, methyltransferases ('writers'), binding proteins ('readers'), and demethylases ('erasers'), all of

which are responsible for dynamic RNA methylation and its functions in posttranscriptional regulation.

m6A writers, erasers, and readers

m6A writers

In addition to the abovementioned METTL3, recent progresses have identified other components of the m6A methyltransferase complex in mammals. In 2014, several independent groups described that the m6A methyltransferase complex contains METTL3, METTL14 and Wilms Tumor 1 Associated Protein (WTAP) as key components (Liu et al. 2014; Ping et al. 2014; Wang et al. 2014b). vir-like m6A methyltransferase associated (KIAA1429), a component of the methyltransferase complex, was found to be required for methylation (Schwartz et al. 2014b). Later, m6A formation in the long noncoding RNA X-inactive specific transcript (*Xist*), as well as in cellular mRNAs, was reported and m6A formation is mediated by RNA-binding motif protein 15 (RBM15) and its paralog RBM15B (Patil et al. 2016). Zinc finger CCCH domain-containing protein 13 (ZC3H13) also regulates nuclear RNA m6A methylation as a novel interactor of m6A methyltransferase complex components (Guo et al. 2018; Knuckles et al. 2018; Wen et al. 2018). The structural basis of the METTL3–METTL14 complex revealed that METTL3 primarily functions as the catalytic core, while METTL14 serves as an RNA-binding platform (Wang et al. 2017). Finally, METTL3 was detected to bind with RNA polymerase II, suggesting that the m6A writer complex is recruited to RNA polymerase II to induce methylation and that this event occurs cotranscriptionally (Slobodin et al. 2017). As this thesis was being written, two studies demonstrated that mechanistic target of rapamycin complex 1 (mTORC1) stimulates m6A mRNA methylation by enhancing WTAP expression and S-adenosine methionine (SAM) synthesis (Cho et al. 2021; Villa et al. 2021). Both studies showed that mTORC1 enhances m6A writer activity upon insulin stimulation (Cho et al. 2021; Villa et al. 2021). However, it is not clear why so many components are recruited to the METTL3–METTL14 complex to perform methylation.

m6A erasers

Fat mass and obesity-associated protein (FTO) and α-ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) are two m6A ‘erasers’ that convert m6A into A. FTO demethylates m6A in both DNA and RNA *in vivo* and that FTO knockdown with siRNA led to increased amounts of m6A in mRNA (Jia et al. 2011). Further discoveries showed that FTO not only removes m6A but also acts on m6Am and reduces the stability of m6Am-containing mRNAs

(Mauer et al. 2017). Wei and colleagues confirmed that FTO binds multiple RNA species and has demethylase activity toward m1A in specific tRNAs, toward m6Am in some snRNAs, and toward internal m6A and cap-m6Am in mRNA (Wei et al. 2018). Collectively, these results suggest that RNA demethylation by FTO is a context-dependent mechanism.

In 2013, ALKBH5 was discovered as the second m6A eraser that oxidatively reverses m6A in mRNA both *in vitro* and *in vivo* (Zheng et al. 2013a). *Alkbh5*-deficient male mice were found to have increased levels of m6A in mRNA and impaired spermatogenesis (Zheng et al. 2013a). ALKBH5 mediated m6A removal in the nuclei of spermatocytes and modulated the splicing and stability of long 3' UTR mRNAs in male germ cells and that failure to do so led to aberrant splicing and production of shorter transcripts (Tang et al. 2018).

m6A readers

m6A affects the fate of mRNAs by recruiting specific 'readers' to m6A-containing mRNA. Several m6A reader proteins have been identified that favor m6A probes in mammalian cells. Five YT521-B homology (YTH) domain-containing proteins were verified to contain m6A-binding domains: YTHDC1, YTHDC2 and the YTHDF1–3 protein family.

YTHDC1, the nuclear member of the highly conserved YTH protein family, not only is an m6A-binding protein but also preferentially recognizes GG(m6A)C sequences (Xu et al. 2014). YTHDC1 has been linked to promoting mRNA splicing (Xiao et al. 2016; Kasowitz et al. 2018), *Xist*-mediated X chromosome silencing (Patil et al. 2016) and facilitating the nuclear export of m6A-modified mRNAs (Roundtree et al. 2017b). Recent studies revealed that YTHDC1 recognizes m6A on chromosome-associated regulatory RNAs in the nucleus and regulates the fate of m6A-modified RNA transcripts to maintain cellular integrity (Chelmicki et al. 2021; Chen et al. 2021a; Liu et al. 2021; Xu et al. 2021).

In 2012, YTHDF2 and YTHDF3 were found to bind exclusively to m6A-modified baits (Dominissini et al. 2012). Later, a study showed that m6A is selectively recognized by YTHDF2 to regulate mRNA degradation (Wang et al. 2014a). YTHDF3 facilitates the translation and decay of m6A-modified RNAs (Shi et al. 2017).

YTHDF1 was initially demonstrated to selectively recognize m6A-modified mRNAs, and YTHDF1 interacts with the translation initiation machinery and enhances the translation efficiency of its target RNAs (Wang et al. 2015).

YTHDC2 is another cytoplasmic m6A reader that regulates the levels of m6A-modified transcripts, which ensures proper meiosis in the testes (Bailey et al. 2017; Hsu et al. 2017; Wojtas et al. 2017). An RNA-activated ATPase fuels the 3′→5′ helicase activity of YTHDC2, suggesting that YTHDC2 promotes mRNA translation (Dhote et al. 2012; Wojtas et al. 2017).

m6A and histone modifications

m6A is deposited on mRNAs that encode histone modifiers and transcription factors to affect transcriptional regulation. Methylation of RNA occurs in a cotranscriptional manner in the cell nucleus. Recently, m6A was found to modulate transcription at the chromatin level.

In tomato fruit ripening, mRNA m6A methylation exhibits dynamic changes similar to DNA methylation and that m6A methylation is associated with decreased expression of the RNA demethylase gene *slalkbh2*, which is regulated by DNA methylation (Zhou et al. 2019b). m6A also preferentially occupies genes with CpG-rich promoters, suggesting that the DNA-methylated region of m6A host genes may be involved in gene expression (Xiao et al. 2019).

Later, Promoter-proximal RNAP II pausing was found to be regulated by m6A (Akhtar et al. 2020).In addition, m6A was found to be cotranscriptionally deposited under the control of H3K36me3 for local histone modification (Huang et al. 2019a). Li and colleagues demonstrated that the methyltransferase complex METTL3-METTL14 regulates H3K9me2 modification, and that this repressive histone mark is specifically removed by the induction of m6A-modified transcripts (Li et al. 2020). Loss of m6A on carRNAs leads to enrichment of the active histone marks H3K4me3 and H3K27ac, which results in open chromatin at carRNA loci (Liu et al. 2020a).

Firstly, our group showed that m6A modification is detectable on the majority of RNA:DNA hybrids to safeguard genomic stability (Abakir et al. 2020). Later, these results have also been validated by other studies, m6A is involved in R-loop formation and promotes chromatin stability (Yang et al. 2019; Duda et al. 2021).

m6A in gametogenesis and embryogenesis

Gametogenesis is the process by which diploid cells undergo mitosis and meiosis to produce haploid cells. Primordial germ cells are the source cells that give rise to gametes—oocytes in females or spermatozoa in males. The process of sperm production is called spermatogenesis, while the production of eggs is called oogenesis. During gametogenesis, histone modifications and the DNA methylation landscape are profoundly remodeled, and these modifications

contribute to the establishment of germ cells and act as regulatory markers during embryogenesis (Hou et al. 2016; Stewart et al. 2016). Several aspects of these epigenetic modifications can be transmitted to offspring (Skvortsova et al. 2018). The process of gametogenesis is also accompanied by orchestrated waves of gene expression, and accumulation of specific RNA transcripts generates discrete transcriptomes for the mature gametes, either eggs or sperm. Recent evidence suggests that RNA modifications in sperm have an essential role in modulating epigenetic memory in the progeny (Chen et al. 2016). Early embryo development is driven by deposited maternal RNAs, and these maternal RNAs are degraded for zygotic genome activation. Only in recent years has the most abundant modification, m6A, been noted for its role in mammalian gametogenesis and embryogenesis (Klungland et al. 2016).

Role of m6A in spermatogenesis

Spermatogenesis is the sequence of events occurring in the seminiferous tubules of the testis, and a diploid spermatogonium (plural: spermatogonia) produces haploid spermatozoa—i.e., sperm. In the first phase, spermatogonia start with mitotic division, divide by mitosis to renew themselves and differentiate to produce primary spermatocytes. Meiosis I is the process by which a primary spermatocyte produces two secondary haploid spermatocytes. These spermatocytes undergo the second meiotic division to produce four round haploid spermatids. After meiosis II, and by the end of this process, spermatocytes become haploid spermatids. Round spermatids undergo remodeling and differentiation into mature spermatozoa.

Single-cell sequencing data revealed that m6A regulators are expressed during spermatogenesis (**Figure 2**) (Green et al. 2018; Wang et al. 2018). Many m6A regulatory enzymes display testis-specific expression patterns during spermatogenesis; based on their RNA expression, m6A regulators are classified into four main expression patterns: high expression only during mitotic stages (*Hnrnpa2b1, Hnrnpc and Eif3a*); high expression only during meiotic stages (*Ythdc1*); high expression during both mitotic and meiotic stages (*Alkbh5, Cbll1, Mettl14, Mettl16, Virma, Ythdc2, Ythdf2 and Zc3h13*); and low expression throughout spermatogenesis (*Fto, Mettl3, Mettl5, Rbm15b, Rbm15, Wtap and Zcchc4*) (Begik et al. 2020). Writer, reader, and eraser m6A regulators (VIRMA, YTHDC2, YTHDF2, ALKBH5, METTL14, and METTL3) are highly expressed in spermatogonia, suggesting important roles for RNA modifications in sperm development and maturation (Begik et al. 2020).

Germ cell-specific inactivation of the m6A RNA methyltransferase METTL3 or METTL14 results in dramatically decreased m6A levels and a reduced number of spermatogonial stem

cells due to a significant decrease in translational efficiency (**Figure 2**) (Lin et al. 2017). Double knockout of METTL3 and METTL14 in advanced germ cells also led to impaired and abnormal spermiogenesis (Lin et al. 2017). METTL3 was essential for male fertility and spermatogenesis in germ cell-specific *Mettl3*-knockout mice and that *Mettl3* ablation severely inhibited spermatogonial differentiation and prevented the initiation of meiosis (Xu et al. 2017). In a zebrafish study, sperm maturation and sperm motility were found to be significantly reduced in *Mettl3* mutant male fish (Xia et al. 2018).

ALKBH5 is a mammalian RNA demethylase that catalyzes the removal of m6A, and m6A removal catalyzed by ALKBH5 was shown to impact mouse spermatogenesis (**Figure 2**) (Zheng et al. 2013b). *Alkbh5*-deficient male mice exhibited increased mRNA m6A levels and impaired fertility, which resulted in apoptosis that affected meiotic metaphase-stage spermatocytes (Zheng et al. 2013b). ALKBH5 is abundant in mitotic and meiotic male germ cells and that ALKBH5-mediated m6A removal in the nuclei of spermatocytes and round spermatids is crucial for correct splicing and stability of mRNAs (Tang et al. 2018). Huang and colleagues found that meclofenamic acid specifically inhibited FTO demethylase activity in spermatogonia; FTO regulates proliferation and cell cycle progression in spermatogonia through its m6A demethylase activity (**Figure 2**) (Huang et al. 2019c).

m6A readers are essential for male fertility in mice (**Figure 2**). The most studied protein is YTHDC2, which was found to be essential for mouse meiosis in the context of spermatogenesis (Bailey et al. 2017; Hsu et al. 2017; Wojtas et al. 2017; Jain et al. 2018). *Ythdc2*-knockout mice are infertile and have significantly smaller testes than *Ythdc2* wild-type mice, and YTHDC2 binds with MEIOC (Meiosis Specific With Coiled-Coil Domain) to regulate the levels of m6A-modified transcripts to ensure proper meiotic progression (Hsu et al. 2017; Jain et al. 2018). YTHDC2 facilitates the transition from mitosis to meiosis in mouse germ cells, and *Ythdc2*-deficient male germ cells can enter meiosis but have an ambiguous identity and fail to express many meiotic markers (Bailey et al. 2017). In addition, studies have suggested that YTHDC2 enhances the translation efficiency of its m6A-modified targets and decreases their mRNA abundance and that YTHDC2 ablation causes upregulation of m6A-enriched transcripts (Hsu et al. 2017; Wojtas et al. 2017). In contrast to YTHDC2, YTHDC1 is expressed in the nucleus of cells, and testes from *Ythdc1*-knockout males were found to completely lack germ cells, including mitotic spermatogonia, and exhibit a sertoli-cell-only phenotype, revealing that YTHDC1 is essential for spermatogonia survival (Kasowitz et al. 2018). Male *Ythdf2*-knockout mice are fertile, with no observed impact on spermatogenesis, suggesting that the YTHDF2-

mediated regulatory m6A RNA pathway is not important for meiosis in mice (Ivanova et al. 2017). RNA-seq data showed that *Ythdf1* and *Ythdf3* are mainly expressed in spermatogonia, while *Ythdf2* is expressed both in spermatogonia and in spermatocytes (Green et al. 2018). Immunostaining of seminiferous tubules also verified that YTHDF1 is expressed in spermatogonia, while YTHDF2 is expressed in a later stage (Lasman et al. 2020c). Triple knockout of *Ythdf1–3* revealed a redundancy in these readers' role during early development, and *Ythdf2* has a dominant role that cannot be compensated by *Ythdf1* or *Ythdf3* (Lasman et al. 2020c). Zebrafish with loss of *Ythdf2* and *Ythdf3* developed healthy testes and were fertile at rates similar to wild-type zebrafish, suggesting that the function of YTHDF1–3 is redundant during zebrafish development (Kontur et al. 2020).

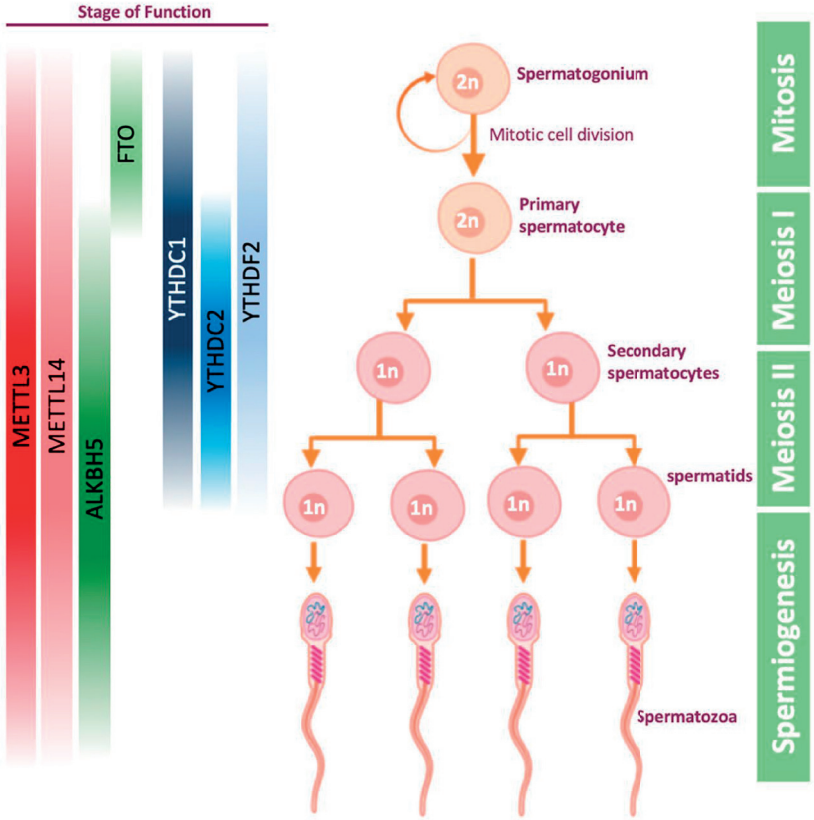


Figure 2. Proteins that have been reported to have roles during spermatogenesis. These proteins have been reported as m6A readers, writers, or erasers in some specific stages. The figure is from (Lasman et al. 2020a).

Role of m6A in oogenesis

Oogenesis is the process by which egg cells are produced in the ovaries. During oogenesis, primordial germ cells undergo a meiotic program to produce functional oocytes. In oogenesis, an oogonium forms during embryological development and mitotically produces approximately 1–2 million primary oocytes by the time of birth. Starting at adolescence, several primary

oocytes enter meiosis I and pause at prophase I in the germinal vesicle (GV) stage. Under the influence of periodic hormone secretion, primary oocytes re-enter meiosis I, producing secondary oocytes and polar bodies. Until fertilization, primary oocytes continue through meiosis II and generate second polar bodies and fertilized eggs—zygotes.

Oocytes are stored with RNA species, which indicates the sustainability of the protein-coding capacity in the early embryo. During the MZT transition, maternal transcripts are cleared from the embryo, whereas zygotic transcripts gradually accumulate. Strong evidence indicates that m6A modification contributes to the destabilizing effect of m6A-modified mRNAs (Zaccara et al. 2019). Considering that maternal mRNA clearance is an essential and dynamic process during MZT and that m6A is the most prevalent internal mRNA modification, it is believed that m6A plays critical roles in maternal RNA degradation (**Figure 3**) (Sha et al. 2019; Vastenhouw et al. 2019).

Early works showed that m6A is essential for proper meiosis in yeast (Schwartz et al. 2013). It is fascinating to explore the effects of m6A on RNA metabolism and its roles in maternal RNA degradation during oogenesis. The first study showed that m6A-dependent RNA decay regulates maternal mRNA clearance during MZT in zebrafish, highlighting the critical role of m6A mRNA methylation in the early stages of embryogenesis (Zhao et al. 2017b). Later, Ivanova et al. revealed that the m6A reader YTHDF2 is essential for the regulation of the maternal transcriptome and oocyte competence in mice (Ivanova et al. 2017). Another study confirmed that *Ythdf1*-knockout and *Ythdf3*-knockout mice are as fertile as their control counterparts, while in female *Ythdf2*-knockout mice, the oocytes show a normal morphology and can be fertilized, but the mice are sterile (Lasman et al. 2020c). Regarding nuclear m6A readers, *Ythdc2*-knockout mice are infertile and have significantly smaller ovaries; in addition, after the mitosis-to-meiosis transition, the oocytes cannot develop beyond the zygotene stage (Bailey et al. 2017; Hsu et al. 2017; Wojtas et al. 2017). *Ythdc1*-deficient oocytes are blocked at the primary follicle stage and lead to massive deficiency in alternative splicing (Kasowitz et al. 2018).

In zebrafish, *Mettl3*-knockout fish are viable, but *Mettl3* knockout leads to failed gamete maturation and to significantly reduced fertility (Xia et al. 2018). In mice, *Mettl3* knockout is embryonic lethal, by utilizing small interfering RNAs or morpholinos to knock down METTL3 in fully grown germinal vesicle (GV) oocytes (Sui et al. 2020). Knocking down METTL3 in female germ cells severely inhibited oocyte maturation and led to defects in MZT (Sui et al. 2020). Conditional *Mettl3* knockout in the ovary was found to have a major effect on oocyte

development and ovarian morphology, and female mice with this genetic alteration were sterile (Lasman et al. 2020c).

Despite these findings, more studies are required to fully understand the mechanisms of oogenesis and early embryo development. The very small amount of RNA isolated from oocytes and early embryos excludes the possibility of performing MeRIP-seq/m6A-seq to determine the topology of the RNA methylomes (Geula et al. 2015b; Ivanova et al. 2017). We developed a method to map the "m6A code" at the ultra-low/single cell level, which could reveal the roles of m6A modification in early mammalian embryogenesis, as presented in this thesis.

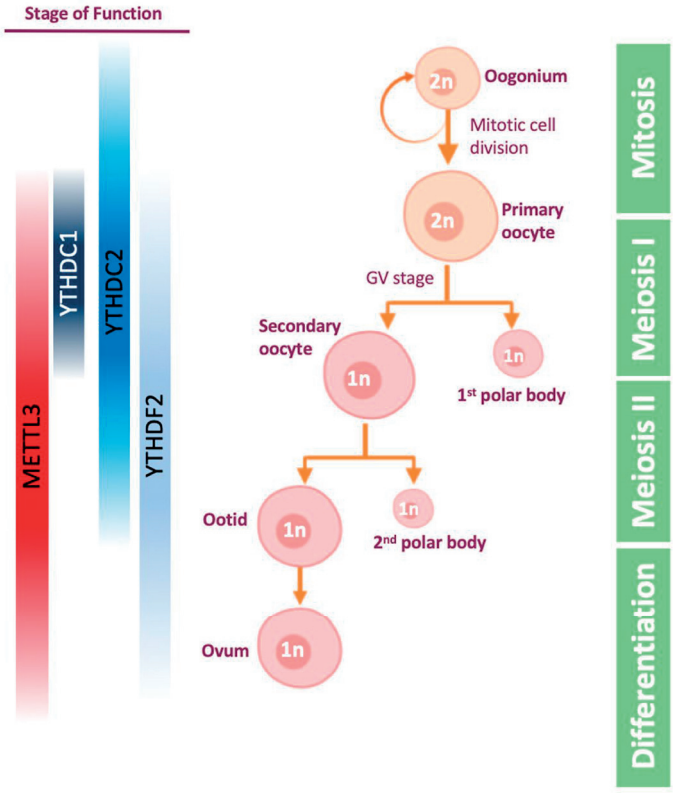


Figure 3. Proteins that have been reported to have roles during oogenesis. These proteins have been reported as m6A readers or writers in some specific stages. The figure is from (Lasman et al. 2020a).

m6A in stem cell and preimplantation embryo development

Embryonic stem cells (ESCs) are derived from the inner cell mass of blastocysts, structures formed during the early stage of preimplantation embryo development. They are able to differentiate into all three primary germ layers: ectoderm, endoderm and mesoderm. Several studies have highlighted that m6A modifications on RNAs play crucial roles in controlling stem cell self-renewal and differentiation (**Figure 4**) (Batista et al. 2014; Wang et al. 2014b; Aguilo et al. 2015; Liu et al. 2020a; Chelmicki et al. 2021; Liu et al. 2021; Xu et al. 2021). In 2014,

Batista et al. found that genetic inactivation or depletion of mouse and human *Mettl3* led to m6A removal from selected target genes, prolonged *Nanog* expression, and impaired ESC exit from self-renewal toward differentiation into several lineages (Batista et al. 2014). Wang Y. et al. demonstrated that knockdown of *Mettl3* and *Mettl14* led to a lack of m6A RNA methylation and loss of ESC self-renewal capability; their knockdown also inhibited the expression of pluripotency genes and promoted the expression of developmental regulators (Wang et al. 2014b). Geula et al. showed that *Mettl3* knockout in naïve embryonic stem cells depleted m6A in mRNAs and failed to adequately terminate the naïve state of these cells, which led to early embryonic lethality (Geula et al. 2015b). Recently, METTL3 was discovered to deposit m6A modifications on carRNAs, that YTHDC1 facilitates the decay of a subset of these m6A-modified carRNAs, and that knockout of the m6A writers *Mettl3* or *Ythdc1* in ESCs increases chromatin accessibility and activates transcription in an m6A-dependent manner (Liu et al. 2020a). In addition, several recent studies identified m6A as an epigenetic mechanism to repress endogenous retroviruses (ERVs): m6A methylation specifically represses ERVs in ESCs to maintain cell integrity and further control histone 3 lysine 9 (H3K9) trimethyltransferase activity, revealing a mechanism of heterochromatin regulation in mammals and highlighting an essential role for m6A in chromatin modification and retrotransposon repression (Chelmicki et al. 2021; Liu et al. 2021; Xu et al. 2021).

As mentioned in the previous section, maternal transcripts are cleared from the embryo, whereas zygotic transcripts gradually accumulate. The mechanism by which m6A is accompanied by the degradation of maternal RNA and is cotranscriptionally deposited on newly synthesized RNA during zygotic genome activation (ZGA) is elusive (**Figure 4**). Notably, immunofluorescence (IF) staining showed that the m6A signal was mainly localized in the cytoplasm during mouse oocyte maturation and embryonic development (Sui et al. 2020). Knocking down *Mettl3* in female germ cells severely inhibited oocyte maturation and led to defects in MZT in preimplantation embryos. Knockdown of hnRNPA2/B1, an RNA-binding protein involved in splicing, increased the m6A intensity, which resulted in delayed embryonic development after the 4-cell stage and blocked further development (Kwon et al. 2019). Loss of YTHDF2 results in failure to regulate the degradation of maternal mRNAs, thus impeding zygotic genome activation in mice (Ivanova et al. 2017). Immunostaining showed that YTHDC1 localized to the nucleus in preimplantation embryos and that loss of YTHDC1 resulted in embryonic lethality (Kasowitz et al. 2018). Recent studies also reported that the

m6A readers IGF2BP1/2 may be involved in the regulation of mRNA stability in ZGA as m6A readers (Liu et al. 2019b; Hao et al. 2020).

However, as mentioned above, it is not feasible to perform MeRIP-seq/m6A-seq to determine the topology of RNA methylomes in preimplantation embryos due to the scarcity of early embryos. Ultra-low input or even single-cell m6A methylome profiling methods are needed to build on current foundations; these approaches would provide insights into how m6A is established in the early embryo.

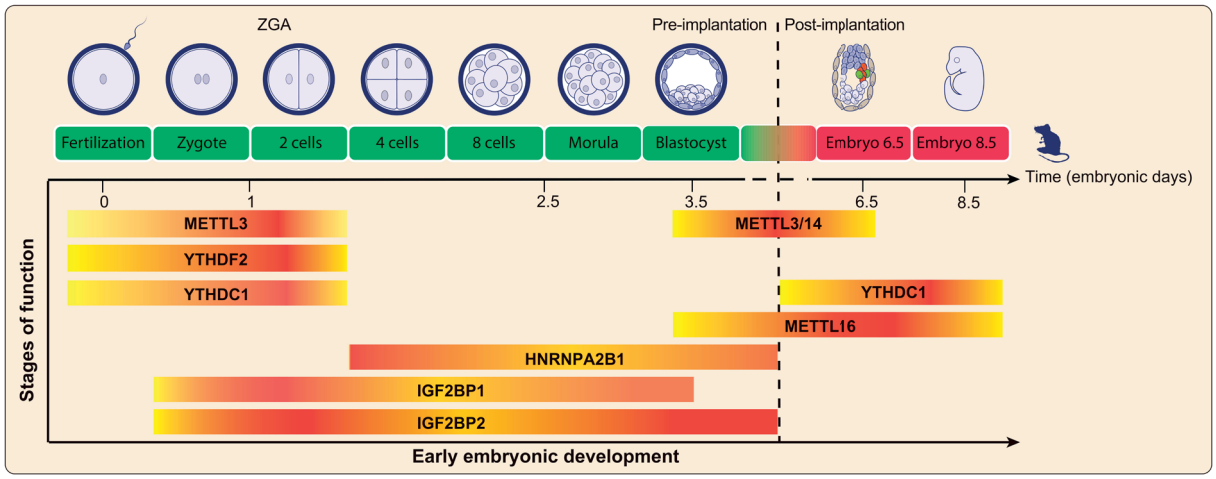


Figure 4. m6A modification-related proteins in embryonic stem cell and preimplantation embryo development. These proteins have been reported as m6A readers, writers, or erasers in some specific stages. The figure is from (Zhang et al. 2020).

m6A detection methods

m6A antibody-based detection methods

The first anti-m6A antibody was characterized in 1977 and was employed to pull down m6A-containing oligonucleotides from tRNA (Munns et al. 1977). m6A-methylated RNAs were first immunoprecipitated (IPed) and sequenced by next-generation sequencing (MeRIP-Seq/m6A-seq) in 2012 (**Figure 5**); this was the first time that m6A on mRNAs was mapped in a transcriptome-wide manner. Polyadenylated RNA isolated from total RNA was fragmented into ~100–200 base pair oligonucleotides, small fractions were aliquoted as the input RNA prior to m6A immunoprecipitation, and the fragmented RNAs were immunoprecipitated using an anti-m6A antibody. The immunoprecipitated RNA and input control fragments were subjected to library preparation and massively parallel sequencing. Reads were uniquely mapped to a reference transcriptome containing a single, intron-free splice variant for each gene. Reads from the IPed samples clustered as distinct peaks, while enrichment of reads in these regions was not

observed in the non-IP control sample. This analysis identified ~12,000 putative m6A sites in the mouse liver and ~7,000 m6A sites in the human brain. Both studies found that the significantly enriched motif was DRACH (D = A, G, or U; R = G or A; H = A, C or U). After that, MeRIP-Seq/m6A-seq methods were widely used to study the role of m6A modification in various diseases. Despite its power, this approach has limitations. First, its resolution is limited by the fragment size of the RNA used for immunoprecipitation—typically 100–200 nt, and it is not feasible to localize the precise m6A sites within the 'peaks'. Second, antibody specificity needs to be assessed before immunoprecipitation, and nonspecific antibodies bring up the issue of false-positive peaks.

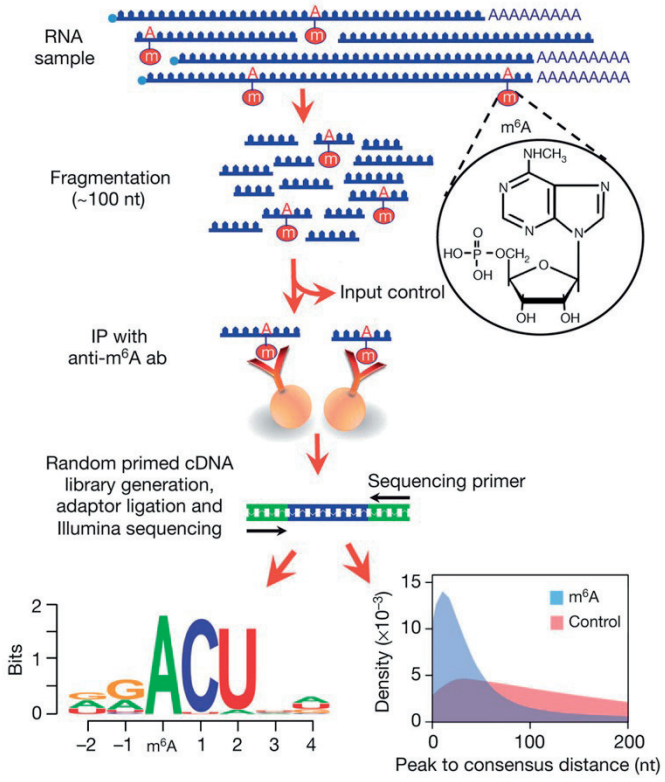


Figure 5. Flowchart of the MeRIP-seq/m6A-seq protocol. ab, antibody; nt, nucleotide. IP, immunoprecipitation. The figure is from (Dominissini et al. 2012). RNAs containing m6A are fragmented and immunoprecipitated by mixing the RNAs with anti-m6A antibody-coupled Dynabeads. m6A-containing RNAs are then eluted from the antibody-coupled beads and subjected to library construction. The resulting highly enriched m6A RNA pool is subjected to next-generation sequencing and mapping. The legend is adapted from (Meyer et al. 2012).

Later, an optimized protocol was reported; this protocol decreased the required mRNA starting material from 400 μg of poly(A)+ mRNA to 5 μg, increased the resolution by decreasing the fragment size to 70~90 nt, employed a ligation-based strand-specific library preparation

protocol capturing both ends of the fragmented RNA, and allowed the removal of false-positives by using strains expressing inactivated m6A writer enzymes as negative controls. This optimized protocol revealed a conserved, widespread, dynamic mRNA methylation program in yeast meiosis at nearly single-nucleotide resolution (Schwartz et al. 2013).

To increase resolution, the strategy of photocrosslinking-assisted m6A-seq was developed (Chen et al. 2015b). 4-Thiouridine (4SU) is incorporated into newly synthesized mRNA in place of U in *in vitro* cell culture, and after the m6A-IP process, the sample is irradiated with 365 nm UV light, which causes 4SU to crosslink with nearby anti-m6A antibodies. The crosslinked RNA is digested into fragments of approximately 30 nt using RNase T1, and the 4SU crosslink is read as a C base in sequencing. This method maps m6A modifications at a resolution of up to 23 nt.

In 2015, two approaches—miCLIP and m6A-CLIP—were developed (Ke et al. 2015; Linder et al. 2015); in these approaches, without 4SU feeding, direct UV-induced RNA-protein crosslinking is performed after IP, and this crosslinking causes unique C-T transition signatures neighboring m6A marks, which are used to map methylation across the transcriptome at single-nucleotide resolution (**Figure 6**).

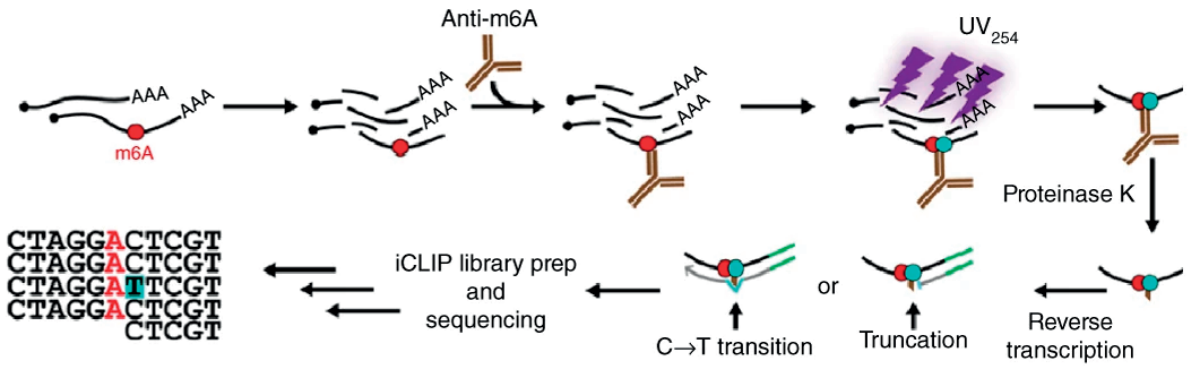


Figure 6. The miCLIP protocol. Purified cellular RNA is fragmented and incubated with anti-m6A antibodies. After crosslinking with UV light (254 nm), covalently bound antibody-RNA complexes are recovered by protein A/G affinity purification and subjected to SDS–PAGE and nitrocellulose membrane transfer. RNA is then released from the membrane by proteinase K and reverse transcribed. Peptide fragments that remain on the RNA lead to nucleotide incorporation errors (indicated as 'C→T transition') and cDNA truncations. The figure and legend are from (Linder et al. 2015).

Although the above antibody-based detection methods provide significant insights into m6A mapping, methods to quantify the methylation level in a stochiometric manner are lacking. In 2016, m6A-level and isoform-characterization sequencing (**Figure 7**) (m6A-LAIC-seq), which globally quantifies the proportion of m6A-modified transcripts, was developed (Molinie et al.

2016b). The protocol is similar to that of m6A-seq, but it omits the fragmentation step and maps full-length transcripts in both the flowthrough (m6A-negative) and the immunoprecipitated (m6A-positive) portions of anti-m6A IP. A spike-in is added as an internal control for strict quantification. The m6A levels in each gene is quantified as the RNA abundance ratio: (eluate) / (eluate + supernatant). It is important to note that this method cannot distinguish the exact methylation status at an internal m6A site at a specific position.

As mentioned above, the major drawback of antibody-based m6A mapping has been the requirement for a large amount of material for effective immunoprecipitation and sequencing. This limitation has hampered application to targets with few cells, such as mammalian oocytes and early embryos, small tissue biopsies, and rare stem cell populations. With the development of advances in library preparation, some studies have reported that optimized low-input m6A-seq protocols can profile the m6A methylome using as little as 500 ng of total RNA (Merkurjev et al. 2018; Zeng et al. 2018).

In our laboratory, we made a substantial breakthrough in antibody-based m6A mapping. In **paper I** and **paper II** of this thesis, we initially describe transcriptome-wide single-cell m6A mapping (we termed this approach scMeRIP-seq). Methylated RNA was immunoprecipitated with an anti-m6A antibody prior to single-tube library construction and sequencing with a high-throughput DNA sequencer. This technique is an easy, robust and reliable technique to map m6A at the single-cell level opening up new avenues for epitranscriptomic study.

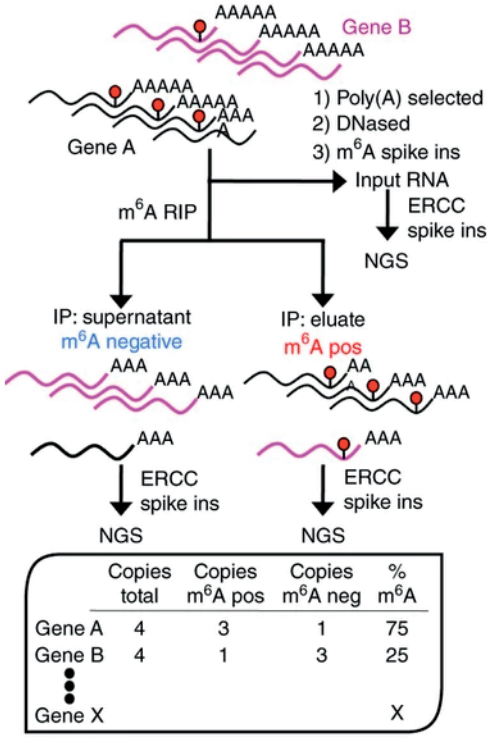


Figure 7. Schematic of the m6A-LAIC-seq protocol. NGS, next-generation sequencing. Pos, positive. Neg, negative. RNA is treated with DNase, and the spike-in controls are added to the poly(A)-selected RNA. Poly(A)-selected RNA containing m6A is immunoprecipitated by mixing the RNA with m6A antibody-coupled Dynabeads directly without RNA fragmentation. After immunoprecipitation (IP), the supernatant fraction is collected. RNAs are then eluted from the antibody-coupled beads. Both the eluted IP and supernatant fractions are subjected to library construction. After next-generation sequencing and mapping, the m6A levels in each gene is quantified as the RNA abundance ratio: (eluate) / (eluate + supernatant). The figure and legend are adapted from (Molinie et al. 2016b).

m6A antibody-free detection methods

The MazF toxin is an ACA-sequence-specific endoribonuclease in *Escherichia coli*. In 2017, a lab found that MazF cleaves RNAs at 5′-ACA-3′ sequences but not at the methylated 5′-(m6A)CA-3′ counterpart sequences (Imanishi et al. 2017). Based on the specific cleavage ability of the MazF RNase, two groups developed MAZTER-seq (Garcia-Campos et al. 2019) and m6A-REF-seq (Zhang et al. 2019b). Both protocols follow common steps: digestion of mRNA with MazF, followed by library preparation and high-throughput sequencing (**Figure**

8). The number of reads at cleavage sites is quantified to calculate the difference between cut and uncut reads. These methods also have limitations; given the requirement of an ACA motif for MazF cleavage, they identify only a subset of m6A sites, which constitutes ~16% of the m6A (only ACA-containing) sites in the mammalian system. These methods do not allow absolute quantification, although the authors claimed quantitative tracking of

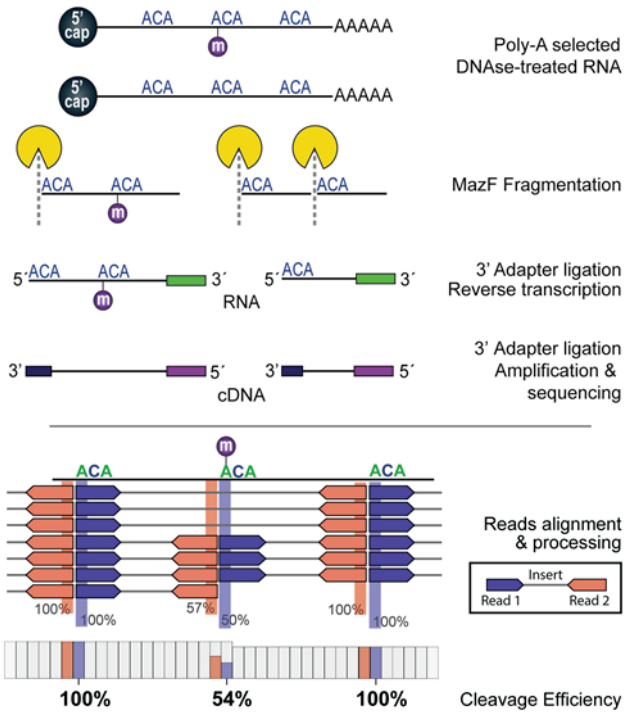


Figure 8. Outline of the MAZTER-seq experimental procedure. (1) poly(A)-selected RNA and DNAse-treated RNA. (2) digestion of mRNA with MazF. (3) end repair, ligation of an adapter to the 3′ end of the resulting RNA fragments, and reverse transcription primed from the ligated adapter. (4) ligation of a second adapter to the 3′ end of the cDNA, cDNA amplification by PCR, and paired-end sequencing. (5) read mapping and identification and quantification of methylation sites. The bacterial RNase MazF cleaves RNA at unmethylated sites at ACA motifs but not at their methylated counterparts. The figure and legend are adapted from (Garcia-Campos et al. 2019).

m6A in diverse biological settings, and their utility depends on the specificity and efficiency of the cleavage enzyme. In addition, the insert size in the sequencing libraries can introduce bias into the results.

In 2020, a metabolic labeling-based m6A detection method called 'm6A-label-seq' was developed (**Figure 9**) (Shu et al. 2020). The N6-methyl group of m6A does not interfere with reverse transcription (RT) when paired to thymine, which is the main obstacle to developing an

RT-based detection method similar to those for other modifications. SAM, a cofactor of methyltransferase, transfers its methyl group to the N6 position of adenines in mRNA. In this study, the authors fed cultured cells Se-allyl-L-selenohomocysteine, which substitutes the normal methyl group of SAM with the allyl group. This allyl-substituted cofactor, catalyzed by the m6A methyltransferase METTL3-METTL14 complex, replaces the original mRNA m6A site to generate N^6-allyladenosine (a6A). Cellular metabolic a6A-labeled RNAs are then enriched with an anti-N6-isopentenyladenosine antibody. These a6A locations are distinguished based on misincorporation during reverse transcription to cDNA.

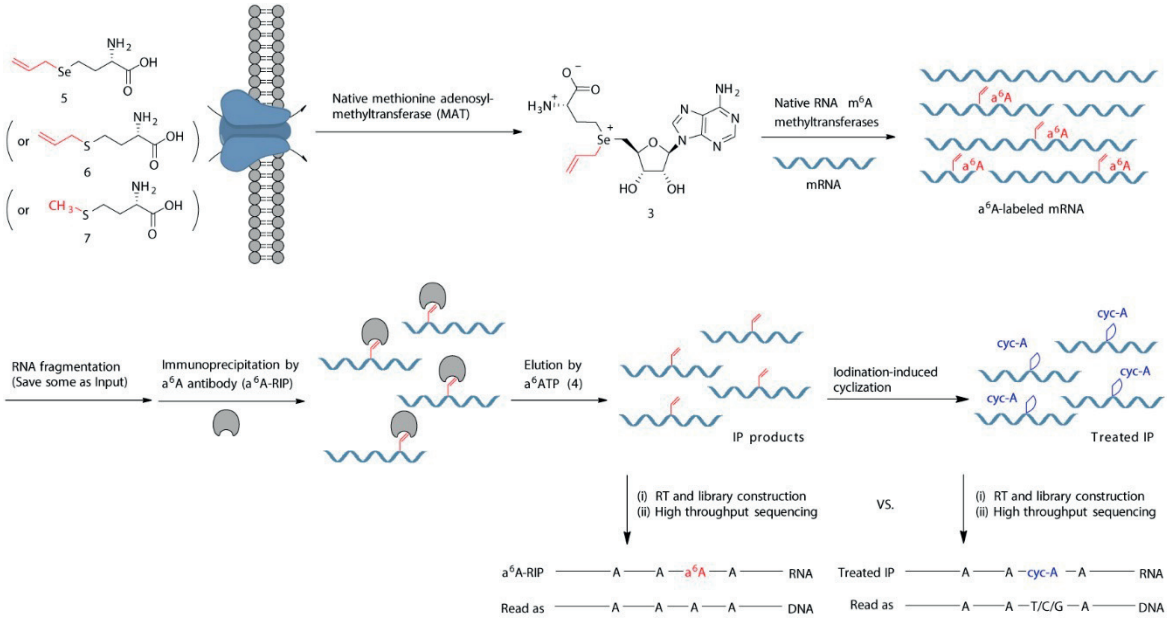


Figure 9. The workflow of m6A-label-seq. Cellular mRNA m6A labeling, RNA enrichment, chemical treatment, cDNA library construction, and high-throughput sequencing. The figure and legend are from (Shu et al. 2020).

At the same time, an FTO-assisted chemical labeling method termed m6A-SEAL-seq was developed for specific m6A detection (**Figure 10**) (Wang et al. 2020). FTO oxidizes m6A to the more stable N^6-dithiolsitolmethyladenosine (dm6A), which enables the conjugation of biotin-labeled tags to dm6A, thereby facilitating affinity purification without the need for an anti-m6A antibody. It is noted that the resolution of m6A-SEAL-seq is comparable to that of MeRIP-Seq/m6A-seq (~200 nt).

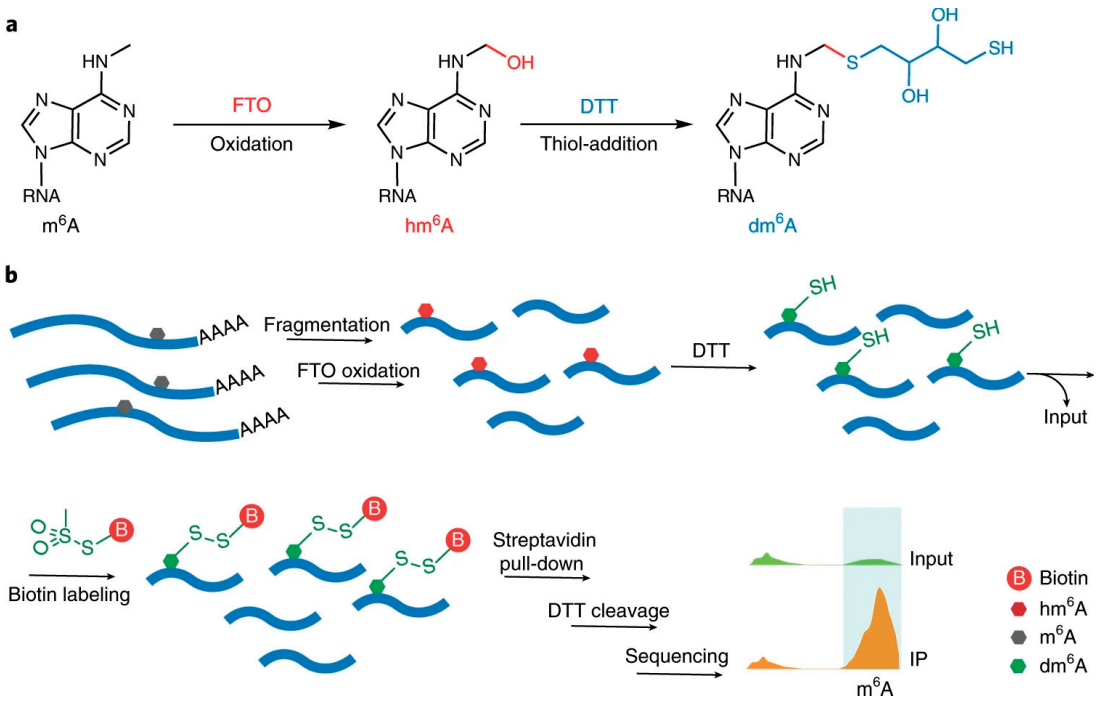


Figure 10. The flowchart of m6A-SEAL-seq. **a,** FTO oxidation of m^6A to hm^6A and the DTT-mediated thiol-addition reaction by which unstable hm^6A is converted to the more stable dm^6A. **b,** Schematic illustration of the m^6A-SEAL-seq pipeline. The figure and legend are from (Wang et al. 2020).

During the writing of this thesis, meCLICK-seq was developed (**Figure 11**) (Mikutis et al. 2020a); in this method, RNA methyltransferase activity is hijacked to introduce a SAM cofactor surrogate (SeAdoYn) to form a propargyl (alkyne-tagged) modification on an RNA substrate in cultured cells. The subsequent copper(I)-catalyzed azide–alkyne cycloaddition reaction with the click degrader leads to RNA cleavage and degradation. After that, RNAs are extracted and sequenced. The harnessing of click chemistry for the degradation of m6A sites provides a clever and simple way to detect m6A in a transcriptome-wide manner. However, only synthetic oligomers are used for *in vitro* studies, lacking some evidence for application in transcriptome-wide detection of m6A in noncultured cells.

Gene editing-based detection methods

APOBEC1 is a zinc-containing cytidine deaminase that induces cytosine-to-uracil (C-to-U) editing (Navaratnam et al. 1993). It has been used in CRISPR–Cas9-based genome editing to induce C-to-U conversion at sites in DNA (Komor et al. 2016). In 2019, a lab used APOBEC1 to edit m6A-adjacent cytidines in RNAs by fusing APOBEC1 to the m6A-binding YTH domain. APOBEC1-YTH expression in cells induces C-to-U deamination at sites adjacent to m6A residues, and total RNA is extracted and subjected to RNA-seq; this method is named DART-seq (**Figure 12**) (Meyer 2019b). The researchers demonstrated that DART-seq can detect m6A

sites in cultured cells from 10 ng of total RNA. However, APOBEC1-YTH overexpression may not be suitable or desirable in some studies, hampering the utility of this method. For example, it cannot be used with materials without editing, although the researchers showed that DART-seq also identified m6A sites *in vitro* but showed reduced efficiency and needed further optimization. The findings also indicated that DART-seq has high false-negative and false-positive rates because of off-target effects (Wang et al. 2020).

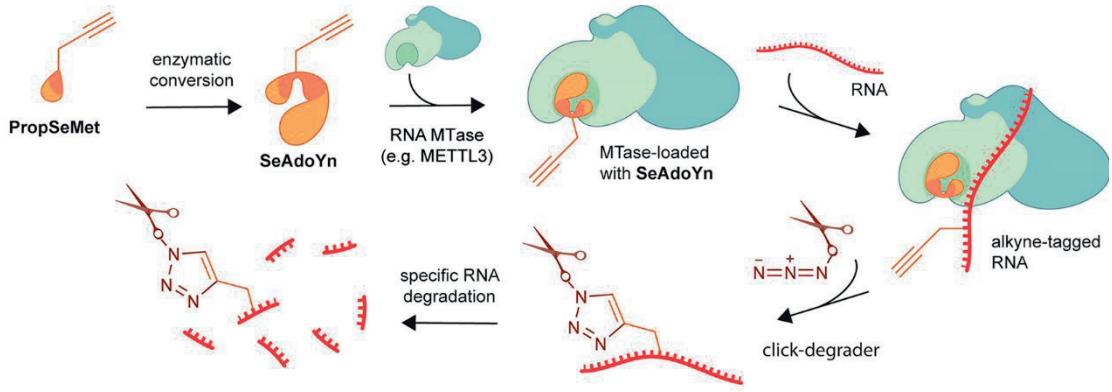


Figure 11. meCLICK-Seq. A small molecule-based method for global detection of RNA methylation by click degradation. meCLICK-Seq involves just two cell culture treatment steps. In the first step, cells are incubated with the methionine surrogate PropSeMet. Cells take up the methionine surrogate, and native methionine adenosyltransferases transform it into SeAdoYn, an S-adenosyl methionine (SAM) surrogate. RNA methyltransferases (MTases), including METTL3 and METTL16, use SeAdoYn as a cofactor instead of SAM, which leads to the addition of propargyl instead of methyl groups onto RNA. A Cu(I)-catalyzed azide–alkyne cycloaddition reaction is then carried out directly in cultured cells to tag RNA with a click degrader, which acts as a functional artificial RNA modification that catalyzes the cleavage of RNA and leads to its degradation. The final step is RNA extraction followed by sequencing. The figure and legend are adapted from (Mikutis et al. 2020b).

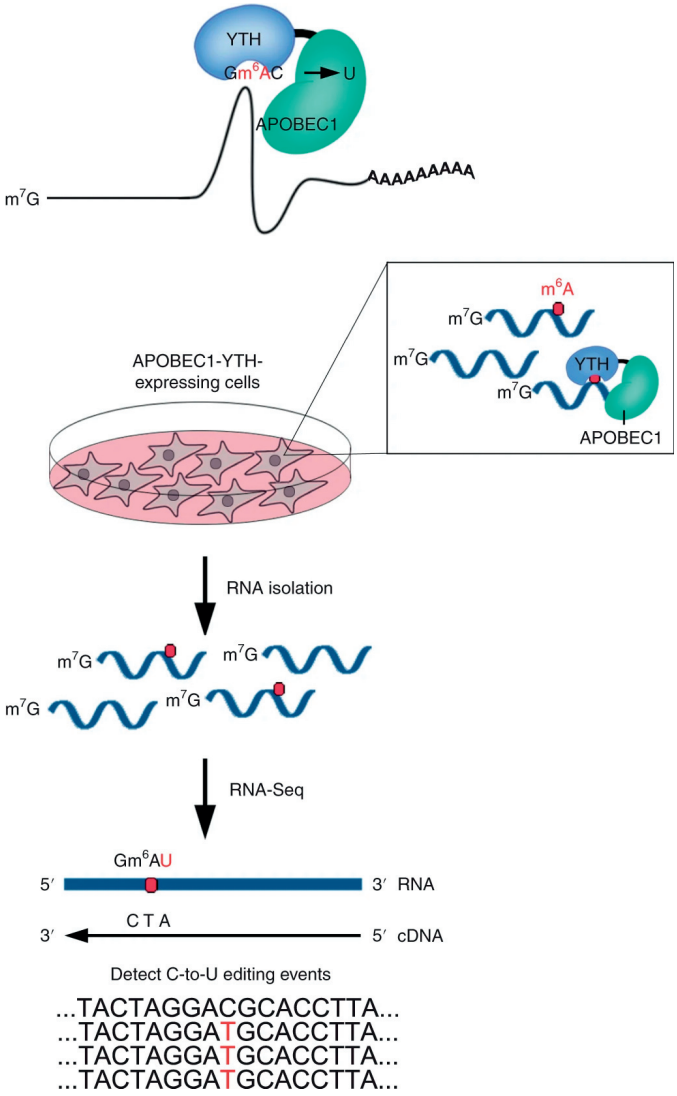


Figure 12. Schematic of the DART-seq method. APOBEC1 is fused to the YTH domain to guide C-to-U editing at cytidine residues adjacent to m6A sites. APOBEC1-YTH is expressed in cells, and total RNA is isolated and subjected to RNA-seq. C-to-U mutations are then detected to identify m6A sites. The figure and legend are from (Meyer 2019b).

Direct RNA detection methods

The current m6A detection methods have many limitations; for example, antibody-based detection methods are highly dependent on antibody specificity, while enzyme-assisted detection methods rely on enzyme efficiency and preference motifs. For library preparation, some methods use high-throughput sequencing with short read lengths, and the inefficient processes of reverse transcription and amplification can introduce biases into the final library. Third-generation sequencing (TGS, also called long-read sequencing) does not require sample amplification and can be used for single-molecule sequencing. There are two commercial third-generation DNA sequencing technologies, Single Molecule Real-Time (SMRT) sequencing (Pacific Biosciences (PacBio)) and Oxford Nanopore Technologies (ONT) sequencing. In

SMRT sequencing, DNA polymerases catalyze the uninterrupted incorporation of fluorescently labeled deoxyribonucleoside triphosphates into the complementary template, light is emitted, and the optical signal is measured in real time until the reads are extended to tens of kilobases. ONT sequencing detects single molecules as they traverse through a nanopore; each nucleoside produces a unique voltage and induces different currents, allowing nucleosides to be distinguished by current alterations.

In 2013, PacBio showed that phospholinked nucleotide binding was affected by m6A in the RNA template and that reverse transcription signals could be monitored in real time by the difference between m6A and A (**Figure 13**) (Vilfan et al. 2013). In ONT sequencing, m6A modifications on RNA cause a detectable current blockade within the nanopore that can be measured (**Figure 14**) (Garalde et al. 2018). Thus, the m6A signal is revealed in the ONT single-molecule reads.

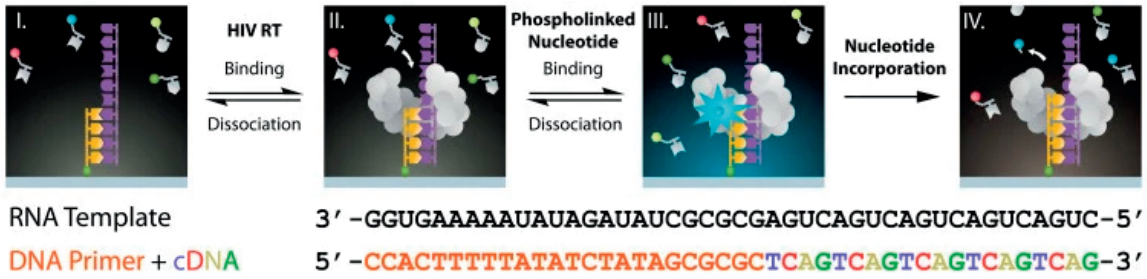


Figure 13. Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. The RNA template (purple) is hybridized to a biotinylated DNA primer (orange) and immobilized at the bottom of the ZMW (I). Reverse transcriptase (RT, gray) binds to the immobilized hybrid (II). Upon initiation of reverse transcription, a correctly base paired phospholinked nucleotide binds in the enzyme's active site (III). The bound nucleotide can either dissociate from the complex (reverse reaction) or become incorporated into the growing DNA chain, accompanied by release of the labeled pyrophosphate (IV). HIV RT then translocates to the next position, and the reaction cycle (II) through (IV) repeats. The figure and legend are from (Vilfan et al. 2013).

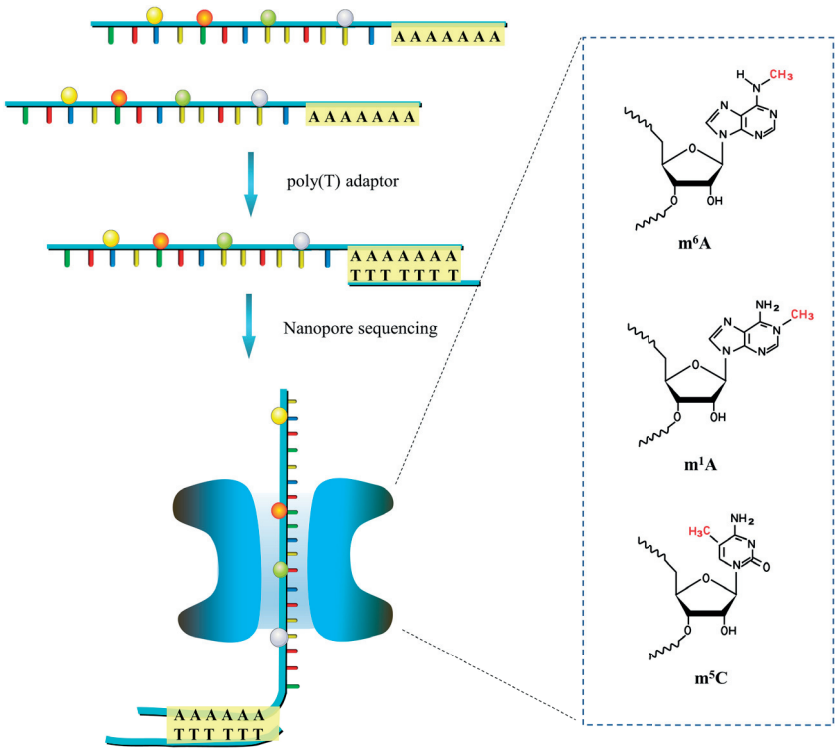


Figure 14. Schematic of direct detection of m6A modification by nanopore sequencing. Oxford Nanopore Technologies' nanopore-based platform detects single molecules of DNA, proteins and small molecules as they traverse through a nanopore. By monitoring current changes, the four biogenic RNA nucleobases and various modified bases, including I, m^6A, and m^5C, can be distinguished. The figure and legend are from (Zheng et al. 2020).

However, neither method is mature or widely used, and there are many unsolved challenges in the application of direct RNA sequencing for m6A detection. For example, the enzymes and chips used in PacBio machines are not commercially available. For ONT sequencing, m6A is not included in the training data for nanopore base-calling software, and an efficient and accurate RNA modification detection algorithm is lacking. A large amount of material is required for input, and when the RNA availability is limited (as it often is), it severely affects the applicability of this technology to rare biological samples. To date, there has been only one study extending nanopore m6A detection to yeast datasets, and the researchers observed that m6A RNA modifications were lacking in yeast ime4-knockout strains (Liu et al. 2019a). More studies are needed to achieve a high accuracy of RNA modification detection with TGS technology and decrease the amount of input material required for RNA modification detection because RNA availability is often limited in some research settings.

Limitation and perspectives

Research into the development of m6A methodology is rapidly progressing, and taking advantage of sequencing techniques has increased our capability to identify m6A in a transcriptome-wide manner. In **Table 1**, the detection methods for m6A modification are compared comprehensively in terms of material requirements, protocol time, sensitivity, etc. Despite these findings, understanding the biological impact of m6A modification still presents several conceptual challenges.

A gold standard mapping method should yield accurate and quantitative information about m6A from the transcriptome. In addition, the whole protocol should be simple and easy for researchers who investigate m6A modification in different cellular and environmental contexts. Moreover, the techniques should be able to be performed with reasonable sequencing costs and effort.

The current m6A detection methods use large cell populations, which hinders their application to rare and precious cell populations. Single-cell m6A mapping is anticipated to reveal not only quantitative expression differences but also dynamics in cell lineage origin and new subtypes. In this thesis, we developed single-cell MeRIP-seq (scMeRIP-seq), which is an antibody-based method for mapping m6A in a transcriptome-wide manner. The material we used was single mouse oocytes, whose RNA content is ten times higher than that of single somatic cells. In the future, it will be possible to optimize the sonication step, use automatic equipment and introduce a combinatorial indexing strategy to achieve a highly scalable assay for m6A methylation profiling of single somatic cells.

In addition, determining how many m6A modifications occur in the same transcripts and quantifying the stoichiometry is always in demand among the epitranscriptomic community. Methods allowing absolute quantification and determination of the correct m6A stoichiometry are anticipated to become pivotal tools for determining the roles of m6A modification in RNAs.

Table 1. The current detection methods for m6A modification

Technique	Starting material	Main technique	Resolution	Time required	Initial quantity	RNA type	Stoichiometric information	Sensitivity	Reference
MeRIP-seq/ m^{6}A-seq	Any sample	IP by m^{6}A antibody NGS	100-200 nt	Long	>5 μg mRNA or >300 μg total RNA	mRNA lncRNA	N/A	Medium	Dominissini et al. (2012); Meyer et al. (2012)
PA-m^{6}A-seq	Cells pre-cultured in 4SU medium	IP by m^{6}A antibody Photo-crosslinking NGS	≥23 nt	Long	10μg ploy (A)+ RNA	ploy (A)+ RNA	N/A	Medium	Chen et al. (2015)
m^{6}A-CLIP/IP	Any sample	IP by m^{6}A antibody Photo-crosslinking NGS	1 nt	Long	3 μg ploy (A)+ RNA	ploy (A)+ RNA	N/A	High	Ke et al. (2015)
miCLIP	Any sample	IP by m^{6}A antibody Photo-crosslinking NGS	1 nt	Long	20 μg ploy (A)+ RNA	ploy (A)+ RNA	N/A	High	Hawley and Jaffrey (2019); Linder et al. (2015)
m^{6}A-LAIC- seq	Any sample	IP by m^{6}A antibody NGS	1500 nt	Long	150μg total RNA	ploy (A)+ RNA	Semi-stoichiometric	Medium	Molinie et al. (2016)
MAZTER-seq	Any sample	MazF digestion RT-qPCR or NGS	1 nt (Only ACA motif)	Medium	100 ng ploy (A)+ RNA	ploy (A)+ RNA	Stoichiometric (Only ACA motif)	High	Garcia-Campos et al. (2019)
m^{6}A-REF-seq	Any sample	MazF digestion FTO digestion NGS	1 nt (Only ACA motif)	Medium	100 ng-200 ng mRNA	mRNA	Stoichiometric (Only ACA motif)	High	Zhang et al. (2019)
DART-seq	APOBEC1-YTH expression cells	Gene editing Cell transfection NGS	10 nt	Long	1 μg total RNA	ploy (A)+ RNA	N/A	Low	Meyer (2019)
m6A-SEAL-seq	Any sample	FTO-assisted chemical labeling NGS	100-200 nt	Long	1 μg poly(A) RNA	mRNA	N/A	Medium	Wang et al. (2020)
m^{6}A-label-seq	Cells pre-cultured in allyl-SeAM medium	Metabolite labeling NGS	1 nt	Long	5 μg a^{6}A-modified	mRNA	N/A	Medium	Shu et al. (2020)
meCLICK-seq	Cells pre-cultured in SAM cofactor surrogate	Metabolite Labeling and cleavage NGS	1 nt	Long	500 ng mRNA	mRNA	N/A	Medium	Mikutis et al. (2020)
SMRT	Synthetic RNA	SMRT	1 nt	Medium	mRNAs N/A	Synthetic	N/A	Low	Vilfan et al. (2013)
Nanopore DRS	Any sample	Nanopore DRS	1 nt	Medium	1 μg ploy (A)+ RNA	RNA ploy (A)+ RNA	N/A	Low	Ayub and Bayley (2012); Liu et al., 2019a; Parker et al. (2020)

Aims of the study

m6A, as the most prevalent and well-characterized internal mRNA modification in higher eukaryotes, has been shown to play crucial roles in RNA splicing, stability, degradation, localization, transport and translation and is involved in spermatogenesis, stem cell differentiation, immune regulation, neurodevelopment, carcinogenesis and other processes (Frye et al. 2018; Zaccara et al. 2019; Barbieri and Kouzarides 2020; Livneh et al. 2020). Since the first method for transcriptome-wide m6A profiling, —MeRIP-seq/m6A-seq (Dominissini et al. 2012; Meyer et al. 2012)—was published in 2012, it has been a powerful technique for transcriptome-wide mapping of m6A sites and has been cited in more than 2000 articles. However, this method requires 300 μg of total RNA, which restricts its application to limited-quantity samples. Moreover, this method requires seven days from initiation to library preparation and next-generation sequencing. These limitations have hindered the application of MeRIP-seq/m6A-seq to small and rare samples, such as cancer biopsies, nonexpandable oocytes and early embryos, and even to the study of heterogeneity in single cells. The studies and results described in this thesis are mainly categorized into three parts:

1. Development of a rapid and efficient ultra-low/single-cell MeRIP-seq method for m6A profiling in a transcriptome-wide manner in single zebrafish zygotes and single mouse metaphase II (MII) oocytes and blastocysts.
2. Presentation of the detailed protocol in a 'recipe' style and step-by-step descriptions of procedures for experimental design, implementation, data analysis and troubleshooting.
3. Application of ultra-low/single cell MeRIP-seq to profile m6A modifications in early mammalian preimplantation embryos.

Summary of papers

Paper I: Single-cell MeRIP-seq maps m6A in oocytes and embryos.

In 2012, the first protocol to allow profiling of transcriptome-wide m6A modification was developed: antibody-based enrichment of m6A-containing RNA fragments followed by massively parallel sequencing. This method has yielded profound insights into the understanding of m6A in cellular functions. However, it still does not meet the demands of the epitranscriptomic community. First, the protocol requires large amounts of RNA samples (≥32 μg of total RNA), which hinders its application to limited-quantity samples. Second, the whole protocol is tedious and time-consuming. These factors have restricted its use in many contexts. In **Paper I**, we developed an ultra-low/single-cell m6A detection method, which opened new opportunities in various studies. For example, it could be applied to analyze cancer biopsies in potential cancer screening, to study mammalian oocyte maturation and preimplantation embryo development in reproduction biology and to explore m6A heterogeneity between single cells. In **paper I**, we developed our highly sensitive single-cell MeRIP-seq (scMeRIP-seq) approach, and we showed that picogram-level m6A detection is achievable, as applied to single zebrafish zygotes as proof of principle. Additionally, by applying scMeRIP-seq to bulk MII oocytes, single MII oocytes and single blastocysts of mice, we found that genes harboring oocyte-specific or blastocyst-specific m6A sites showed distinct relevant functional enrichment patterns in mouse oocytes and blastocysts. Furthermore, we demonstrated that scMeRIP-seq is a powerful approach to not only reveal m6A across mRNA transcripts in single oocytes but also identify m6A in noncoding RNAs. We first found that m6A was predominantly deposited on the MTA_Mm subfamily of mammalian apparent LTR retrotransposons (MaLRs), providing uncharted territory for research on the MZT process.

Paper II:

Transcriptome-wide mapping of N^6-methyladenosine via antibody-based immunoprecipitation and high-throughput sequencing at ultra-low/single-cell resolution.

In **Paper II**, we described the protocol of ultra-low/single-cell MeRIP-seq, suitable for profiling limited-quantity cells, for benchtop application in a step-by-step format. Ultra-low/single-cell MeRIP-seq enables users to measure the transcriptome-wide m6A distribution in any cell type for which only limited material can be obtained. More importantly, this method allows us to understand methylomic heterogeneity in the m6A distribution in single cells (i.e., single MII oocytes in **Paper II**). Finally, our library preparation protocol was specifically designed to retain strand-of-origin m6A information, which allowed us to map reads accurately to their strand of origin.

In summary, methylated RNA is immunoprecipitated with an anti-m6A antibody prior to single-tube library construction and sequencing with a high-throughput DNA sequencer, constituting an easy, robust and reliable technique for mapping m6A at the single-cell level. The whole protocol can be completed within ~4 d from sample collection to data analysis. In brief, the single-cell MeRIP-seq protocol enables m6A analysis at the single-cell level, which helps to determine the m6A methylome in rare materials.

Paper III:

The landscape of m6A in mammalian preimplantation embryos

Currently, there is no work providing information on the m6A landscape of mammalian oocytes and preimplantation embryos. In **Paper I** and **Paper II**, we developed an ultra-low/single-cell MeRIP-seq method suitable for the study of even single oocytes. In **Paper III**, we apply our low-input methyl RNA immunoprecipitation and sequencing (picoMeRIP-seq) method to map m6A in the mouse oocyte and preimplantation embryo transcriptomes. We show that the m6A landscape is highly dynamic during the maternal-to-zygotic transition. Changes in the m6A marking are particularly prominent on genes essential for cell fate determination. The vast majority of stage specifically expressed transcription factors are marked by m6A, and they are more frequently marked by m6A than other transcripts. Maternally-inherited transcripts marked by m6A and degraded post fertilization are more likely to be targeted by miRNA as compared to zygotic transcripts marked by m6A. Moreover, RNAs derived from retrotransposons, such as MTA and MERVL, are marked by m6A in a stage specific manner. Our results provide a foundation for future studies exploring the regulatory roles of m6A in mammalian early embryo development.

Discussion

In this thesis, we described the development, validation, and applications of ultra-low/single-cell MeRIP-seq for both small cell numbers and single-cell analysis. This method is specifically designed for transcriptome-wide mapping of m6A modifications in oocytes and preimplantation embryos. To develop ultra-low/single-cell MeRIP-seq, we rigorously designed several critical experimental steps, including ribosomal RNA (rRNA) depletion, RNA sonication, antibody selection, washing buffer optimization and library construction. We determined the sensitivity and quantitative accuracy of ultra-low/single-cell MeRIP-seq for m6A methylome determination by evaluating it in a mouse liver poly(A)-selected RNA dilution series. To assess transcriptome-wide performance, we generated m6A methylome profiles from 10 ng, 1 ng and 100 pg of mouse liver poly(A)+ RNA, with two technical replicates per RNA amount. We found that even when we used an RNA input at the picogram level, we obtained excellent performance when compared with that observed in previous studies and with our high-level input, which encouraged us to apply this method to single zebrafish zygotes. The overall m6A profiles of single zebrafish zygotes were also excellently consistent with those of bulk zygotes. In addition, by applying ultra-low/single-cell MeRIP-seq to single MII oocytes and single blastocysts of mice, we identified distinct m6A patterns and deconvoluted stage-specific RNA m6A methylation profiles in mouse oocytes and blastocysts (**Paper I**). Next, we described this ultra-low/single-cell MeRIP-seq method, which is an antibody-based method for transcriptome-wide mapping of m6A in a step-by-step manner (**Paper II**). Methylated RNA is immunoprecipitated with an anti-m6A antibody prior to single-tube library construction and sequencing by a high-throughput DNA sequencer, constituting an easy, robust and reliable technique for mapping m6A at the single-cell level. The whole protocol can be completed within ~4 d from sample collection to data analysis. A major advantage of ultra-low/single-cell MeRIP-seq is that the method drastically reduces the amount of material required for m6A profiling. Hence, ultra-low/single-cell MeRIP-seq enabled us, for the first time, to reveal m6A reprogramming during the oocyte-to-zygote transition and preimplantation embryo development in mice (**Paper III**). The IF staining results revealed that high levels of m6A existed in the oocytes and zygotes and that the signals decreased rapidly in 2-cell embryos and gradually increased from the 2-cell stage to the blastocyst stage. The methylome results showed that RNA m6A methylation is controlled in a spatiotemporal manner and participates in the regulation of preimplantation embryo development. Collectively, the results (i) highlight the utility of ultra-low/single-cell MeRIP-seq for the analysis of rare cell populations and single

cells and (ii) provide new insights into the role of m6A during the oocyte-to-zygote transition and preimplantation embryo development. The other detailed technical aspects of ultra-low/single-cell MeRIP-seq and novel biological aspects of the work are discussed below.

Technical aspects of the development of ultra-low/single-cell MeRIP-seq

There are two drawbacks to MeRIP-seq in its application to rare cell samples and single cells. First, the MeRIP-seq protocol needs at least 9 days for completion from beginning to end. Due to the lengthy procedures, sample loss is an unavoidable problem for MeRIP-seq, which hinders its application to low cell numbers and even results in bias in the final library. For immunoprecipitation from large samples, the basic form is very simple—just conjugate an anti-m6A antibody to beads, add the beads to the sample, and pull down the target m6A-modified RNA fragments. However, in MeRIP-seq with single cells or limited samples, there are many factors that need to be optimized (**Paper I** & **Paper II**). Next, the conventional preparation of RNA libraries requires conversion of the RNA input material into cDNA and utilizes ligation of either RNA or DNA adapters to the target RNA or DNA molecules. These adapters are compatible for platform-specific sequencing (Illumina sequencing platforms). Ligation of adapters is not only time-consuming but also inefficient and requires a large amount of starting material. Taken together, these hurdles become the main problems when applying this method to low cell numbers and even single cells.

Signal-to-noise ratio improvement

The scMeRIP-seq approach is similar to chromatin immunoprecipitation sequencing (ChIP-seq); from a bioinformatic analysis perspective, scMeRIP-seq is also based on global identification of significant peaks. In addition to the ability to acquire input material without immunoprecipitation, enriched regions across the genome or transcriptome are identified with an algorithm. In ChIP-seq, a key consideration and challenge is the signal-to-noise ratio; an effective signal-to-noise ratio to separate the signal from the background after immunoprecipitation, library preparation and sequencing is required. Hence, we selected one positive locus (m6A peak) and one negative locus (no m6A peak) from the initial m6A-seq paper (Dominissini et al. 2012), and we used qPCR to quantify the signals from the IP and supernatant fractions, as follows: signal-to-noise ratio = positive locus (m6A peak)/negative locus (no m6A peak). This method was used in our ChIP-seq experiments to optimize several factors in single-cell ChIP-seq.

First and foremost, in ultra-low/single-cell MeRIP-seq, capture of environmental contaminants is very possible. Environmental contaminants such as cellular, bacterial, fungal, or viral RNA can be captured due to the nonsterile environment of the working bench. For this reason, extreme care should be taken to minimize environmental contaminants (e.g., from the air, the bench, pipettes, tube racks, lab coat sleeves, etc.) and avoid the introduction of contaminants into reagents from the kit. The following guidelines should be followed to minimize contamination. First, all experiments were relocated from an open lab bench to a laminar flow (LAF) cabinet, and the bench and equipment were treated with UV irradiation for sterilization. Second, we used cold blocks instead of ice, which could possibly introduce contamination, and all powdered reagents and liquid reagents were prepared in the LAF cabinet. These measures were critical to successfully perform ultra-low/single-cell MeRIP-seq without contamination. In addition, we wore gloves and sleeve covers throughout the whole procedure to protect RNA samples from degradation by any introduced contaminants and nucleases.

For antibody-based immunoprecipitation, first, the antibody was incubated with Dynabeads Protein A in a tube, and the antibody-coated beads were used for m6A target selection. We optimized the antibody/bead ratio to avoid the extra surface area accessible for nonspecific target binding without significant loss of the m6A target. Furthermore, the previous IP procedure (incubation of antibody-coated beads and m6A-containing RNAs) was performed in less than 60 minutes, and we increased the incubation time of the antibody-coated beads with the m6A target to 120 minutes, which yielded increased target capture.

To further reduce nonspecific binding of RNA fragments to the antibody-coated beads, we used the same stringency of washing steps that was used in our single-cell ChIP assays. These washing buffers contained higher detergent (SDS) and salt (NaCl) concentrations, displaying better signal-to-noise ratios not only in the 100 ng input but also in the 10 ng input. The traditional way to carry out washing steps for m6A analysis is to spin the tube on a rotor for head-over-tail rotation. This method is time-consuming and inefficient. We changed this traditional method to rough vortexing at full speed, with four times start/stop cycles to result in acceleration and deceleration of the tube, which ensured improved exchange of buffer around the beads and strong currents to remove the nonspecifically bound target. In our test, we found that we could significantly increase the signal-to-noise ratio by altering the washing method in both the 100 ng and 10 ng inputs. Compared with the traditional head-over-tail rotation method, our method reduced the time for four washes to less than 10 minutes. In addition, the washing

volume is another parameter that needs to be considered in the workflow. In our test, we found that the volume of the washing buffer did not significantly impact the signal-to-noise ratio.

With advances in tube manufacturing technologies, top-quality polypropylene plastic ensures nearly 100% recovery of RNA molecules without the need for a surface coating on the tube to eliminate the risk of sample loss and contamination. Many studies have shown that low-binding tubes help minimize sample loss during immunoprecipitation with lengthy processing steps. For larger inputs, it is possible to perform experiments without using low-binding tubes; however, for ultra-low inputs, it is critical to recover all m6A targets into the final library. We found that the use of Axygen® MAXYMum Recovery® low-binding tubes significantly improved the signal-to-noise ratio, as the product information claimed, likely by reducing loss of RNA to the plastic surface and reducing carryover of nonspecifically bound RNA.

Currently, several polyclonal and monoclonal antibodies that identify m6A have been reported to specifically target m6A-modified RNA. The success of ultra-low/single-cell MeRIP-seq depends on the integrity of highly specific antibodies against m6A; however, the specificity of m6A-specific antibodies has posed a challenge. Some antibodies are nonspecific due to their promiscuous detection of m6A sites. We selected three positive loci (m6A peaks) and three negative loci (no m6A peaks) from the initial m6A-seq paper. The signal-to-noise ratio showed that the Millipore antibody was consistent with the report in the first m6A-seq paper and was superior to the other anti-m6A antibodies in terms of the signal-to-noise ratio.

We failed to recover any RNA fragments with the previous metal ion–induced fragmentation approach, and we reasoned that the time period for chemical fragmentation is difficult to control, especially for ultra-low/single cell inputs. Based on our previous single-cell ChIP-seq experiments, we optimized the sonication conditions using a probe sonicator to obtain the desired fragment sizes. During sonication, RNA samples are subjected to hydrodynamic shearing by exposure to brief periods of sonication. The periods of time are easy to control, allowing the production of an ideal RNA fragment size of approximately 100–500 bp, and it is also tempting to use sonication to avoid overdigestion when compared with chemical fragmentation. It should be noted that the RNA postfragmentation size distribution can only be validated by measuring the DNA concentration after library construction, and it is impossible to measure for ultra-low/single-cell inputs because this low concentration is undetectable by any instrument. The final library should produce a distribution of DNA fragment sizes centered on 200∼600 nt (library preparation adds 139 bp to the size of the fragmented RNA molecules).

single-tube rRNA & DNA depletion

rRNAs are extremely abundant, accounting for more than 90% of the total RNA in mammalian cells. Efficient rRNA removal is important for cost-effective sequencing of RNA samples. After we optimized the above parameters with mouse poly(A)-selected RNA, we further aimed to combine current rRNA removal technology with scMeRIP-seq. The currently available commercial kits for rRNA removal are classified into three distinct types based on removal strategy. First, rRNA is captured by complementary oligonucleotides that are linked to paramagnetic beads, after which the bound rRNA is precipitated and removed from the reaction. Several biotechnology companies provide this type of kit, for example, Illumina's RiboZero, Qiagen's GeneRead rRNA Depletion, and Lexogen's RiboCop kits. In addition, enrichment of poly(A) transcripts can be accomplished by strategies using poly(T) paramagnetic beads. Second, complementary oligonucleotides can also be used with DNA, with hybridization of rRNA to DNA oligos and subsequent degradation of the RNA:DNA hybrids with ribonuclease H (RNase H). These kits can be obtained from several companies, for example, NEBNext rRNA depletion, Kapa RiboErase, and Takara/Clontech's RiboGone. Third, for some very low input amounts, it is challenging to remove rRNA from the initial total RNA. Takara provides an innovative technology allowing the removal of ribosomal cDNA (cDNA fragments originating from rRNA molecules) after cDNA synthesis using probes specific for mammalian rRNA. This strategy allows ribosomal cDNA to be removed after library amplification, which makes it especially well suited for working with very small quantities of total RNA or very small numbers of cells, even single cells. For DNA removal, DNase treatment is currently the best way to remove contaminating DNA from RNA samples. This is a simple and safe method for eliminating DNA contamination from RNA samples without organic extraction or heat inactivation. We wanted to minimize RNA loss during procedures for the removal of both rRNA and DNA in living cells; thus, we reviewed all the available commercial kits that include rRNA and DNA depletion protocols that can be used with intact cells. We found that the cutting-edge NEBNext® rRNA Depletion Kit, which is compatible with a broad range of input amounts from 10 ng–1 μg, is the most promising kit. The kit allows depletion of both cytoplasmic (5S rRNA, 5.8S rRNA, 18S rRNA and 28S rRNA) and mitochondrial (12S rRNA and 16S rRNA) ribosomal RNA from total RNA preparations. The whole protocol for rRNA and DNA depletion was performed in a single tube. We performed tests with zebrafish single zygotes, single MII oocytes and single blastocysts and found that the kit was able to efficiently remove rRNA, resulting in less than 1% rRNA contamination in the final libraries. We are the first to

prove that this kit can be applied to single cells. Finally, as a bonus, this kit uses Agencourt RNAClean XP beads to purify the final RNA while removing all protein contamination.

Other factors in single-cell MeRIP-seq

Antibodies are noncovalently coupled to beads using Dynabeads Protein A. The incubation time for antibody binding to magnetic beads is a key factor in developing scMeRIP-seq protocols. Long incubation times are often used to guarantee the binding efficiency. In our test, we performed antibody-bead incubation overnight. By comparing the final signal-to-noise ratios, we did not find that the binding efficiency was increased at longer incubation times, which means that extending the incubation time longer than overnight did not result in further gain. Considering the temperature sensitivity of the antibody, it should be mentioned that antibody conjugation to the beads should be performed at 4 °C instead of room temperature.

During antibody-bead incubation, the presence of too many beads can influence the background level and result in a low signal-to-noise ratio. We validated various antibody-bead ratios and found that the best ratio was 6.7 μg of antibody/mg of beads for the scMeRIP-seq protocol.

After antibody-bead incubation, the antibody-conjugated beads need to be washed, and the washing conditions are another factor that needs to be considered. To avoid loss of antibodies and ensure the highest yield, we selected a gentle wash buffer, i.e., IP buffer, which will be used in the next step IP and to avoid introducing other unwanted impurities and reagents.

In immunoprecipitation, a small volume keeps the m6A fragment concentration high and therefore increases the binding affinity, which is particularly important for single-cell and rare cell inputs. Previously, MeRIP-seq/m6A-seq required a volume of 500 μl or 1000 μl of reaction mixture during the IP step. We decreased the incubation volume to 100 μl to allow our m6A fragment concentration to remain high and therefore increase the binding affinity.

In the elution step, a previous method used an m6A competitive elution strategy to elute m6A targets; however, the m6A competitive elution buffer for each pulldown is not efficient for low input material. We employed Proteinase K for digestion of the antibodies that connected the beads to the target m6A fragments. Two rounds of Proteinase K digestion were performed to maximize recovery. The recovered eluates from the same sample were pooled and precipitated with ethanol. Ethanol precipitation is an operationally simple, amplification facile, and solvent safe method for extracting RNA. There was also a minor difference from ethanol precipitation in previous large-scale MeRIP-seq/m6A-seq protocols: after salt and ethanol were added to the

aqueous solution, we also added linear acrylamide, which facilitates precipitation in small amounts of RNA. It increases the pellet mass and forces the precipitation of m6A fragments out of the solution. After precipitation, the m6A fragments were separated from the rest of the solution by centrifugation. We washed the pellets twice with 1 ml of ice-cold 75% ethanol to thoroughly remove unwanted contamination. Finally, the pellets were resuspended in nuclease-free water containing an RNase inhibitor.

Preparation of single-cell MeRIP-seq libraries

The current MeRIP-seq/m6A-seq method depends on the construction of DNA libraries for sequencing on an Illumina-compatible platform. The library construction protocol included reverse transcription of RNA to generate complementary single-stranded cDNA, second-strand cDNA synthesis and adapter addition. Adapters can be added either during cDNA synthesis or during the final amplification step. This library preparation protocol has two major drawbacks: first, it requires large amounts of RNA starting material; second, the whole procedure is time-consuming; and third, it is labor intensive and inefficient. We tested several traditional commercial methods for scMeRIP-seq library construction. We found that the ligation of adapters is time-consuming and inefficient; in our hands, we increased the starting amount of input material to at least 10 ng for successful library construction. In addition, the final libraries were contaminated with cross- and self-ligated adapter byproducts that are difficult to remove before sequencing.

Encouragingly, a single-tube DNA library preparation (Carøe et al. 2018) protocol eliminates inter-reaction purification by replacement of the column-based/bead-based purification method with heat inactivation of enzymes. This protocol increases library complexity and yields more uniquely deduplicated reads by reducing sample loss without purification steps. In addition, it is time-saving compared with traditional methods and requires fewer manual manipulations. It also reduces costs when considering the reduced number of steps and decreased laboratory equipment requirements.

In 1999, a method was described for amplifying cDNA ends; it requires only first-strand cDNA synthesis and a single PCR amplification step to generate a final product with very low or nearly no background (Matz et al. 1999). The method was commercialized and is currently used in the Takara SMART-Seq Stranded Kit. It takes advantage of Moloney murine leukemia virus reverse transcriptase (MMLV-RT) to attach adapters to the 5′-ends of cDNAs generated from first-strand synthesis. As described in the user manual of the Takara SMART-Seq Stranded Kit

(**Figure 15**), when the MMLV-RT reaches the 5’ end of the RNA fragment, the enzyme’s terminal transferase activity adds a few nontemplated nucleotides to the 3’ end of the cDNA. The carefully designed SMART-Seq Stranded Adapter undergoes base pairing with the nontemplated nucleotide stretch, creating an extended template to enable MMLV-RT to continue replication to the end of the oligonucleotide. In the next step, full-length Illumina adapters including barcodes are added to the cDNA by PCR amplification.

The SMART-Seq Stranded Kit offers workflow choices between ultra-low (1–50 cells or 10–500 pg total RNA) and low (50–1000 cells or 0.5–10 ng total RNA) input, which meets the requirement for our single-cell MeRIP-seq library preparation. Although this kit is very promising, it still cannot be used directly with our IPed samples. As shown in **Figure 15**, the workflow used in this kit starts with intact cells or total RNA, and removal of rRNA is performed after cDNA synthesis and amplification using probes specific for rRNA.

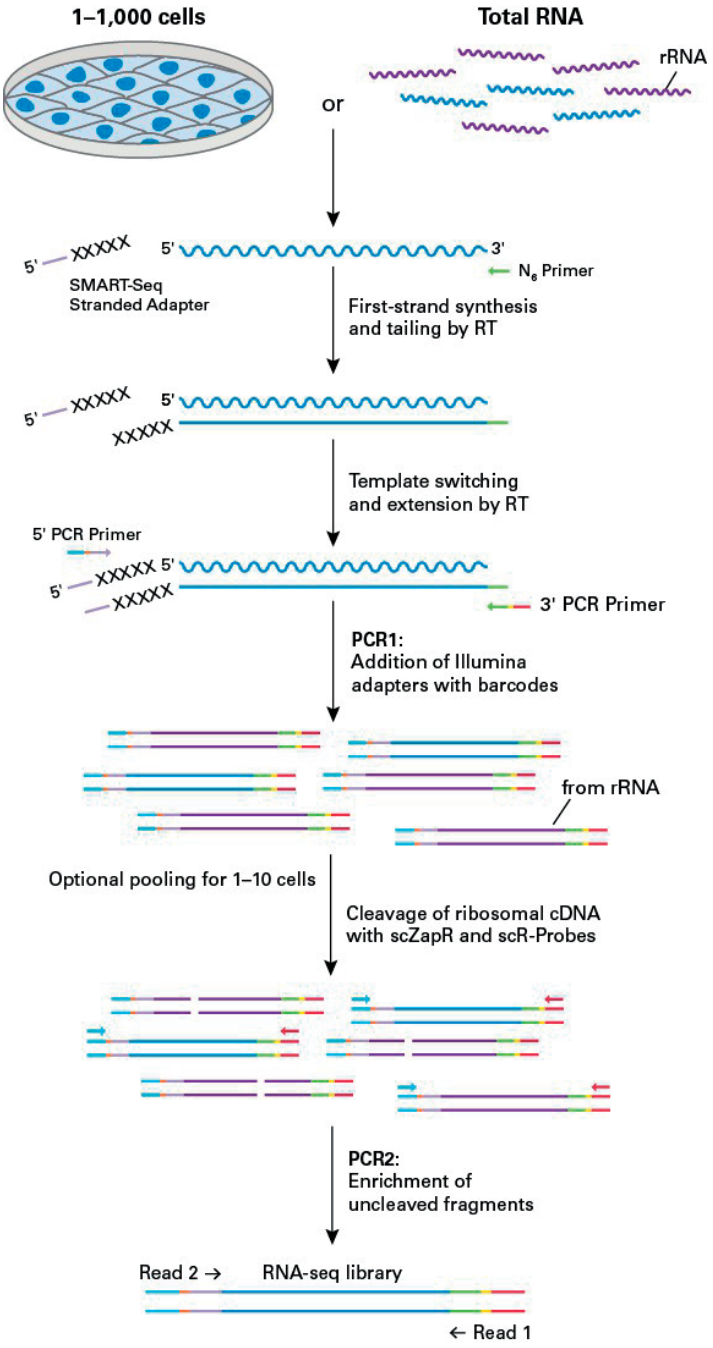


Figure 15. Schematic of the technology used in the SMART-Seq Stranded Kit. SMART technology is used in a ligation-free protocol to preserve strand-of-origin information. Random priming (through the SMART scN6 Primer) allows the generation of cDNA from all RNA fragments in the sample, including rRNA. When the SMARTScribe™ Reverse Transcriptase (RT) reaches the 5' end of the RNA fragment, the enzyme's terminal transferase activity adds a few nontemplated nucleotides to the 3' end of the cDNA (shown as "XXXXX"). The carefully designed SMART-Seq Stranded Adapter (included in the SMART scTSO Mix) undergoes base pairing with the nontemplated nucleotide stretch, creating an extended template to enable the RT to continue replication to the end of the oligonucleotide. The resulting cDNA contains sequences derived from the SMART scN6 Primer and the SMART-Seq Stranded Adapter. In the next step, full-length Illumina adapters including barcodes are added via a first round of PCR amplification (PCR1). The 5' PCR Primer binds to the SMART-Seq Stranded Adapter

sequence (light purple), while the 3' PCR Primer binds to the sequence associated with the SMART scN6 sequence (green). Ribosomal cDNA (originating from rRNA) is then cleaved by scZapR in the presence of mammalian-specific scR-Probes. This process leaves the library fragments originating from non-rRNA molecules untouched, with priming sites available on both the 5' and 3' ends for further PCR amplification. These fragments are enriched via a second round of PCR amplification (PCR2) using primers universal to all libraries. The final library contains sequences allowing clustering on any Illumina flow cell. Optional pooling of up to 12 samples after PCR1 allows for greater ease of use by minimizing the number of samples to be processed downstream. The figure and legends are adapted from the Takara user manual.

However, m6A modifications on rRNA are well known: one is m6A modification on site 4220 in 28S rRNA, and the other is m6A modification on site 1832 in 18S rRNA. We utilized single-tube rRNA and DNA depletion at the beginning of the scMeRIP-seq experiment; initial rRNA depletion from the starting cell lysate leaves a sufficient amount for immunoprecipitation of m6A-modified mRNAs and noncoding RNAs. In addition, initial rRNA depletion from the lysate of single cells is beneficial for reducing the sequencing cost and simplifying analysis of the final results. Then, we contacted Takara's technical support team and consulted with them regarding the detailed protocol. They helped us to design the entire flowchart for library preparation, as shown in **Figure 16**. The modified protocol was very successful in all of our sample library preparations. An added bonus was that it also saved some time by not including rRNA depletion after the first round of library amplification. The entire library construction protocol can be completed in less than 5 hours, while at least 7 hours are needed for the original protocol.

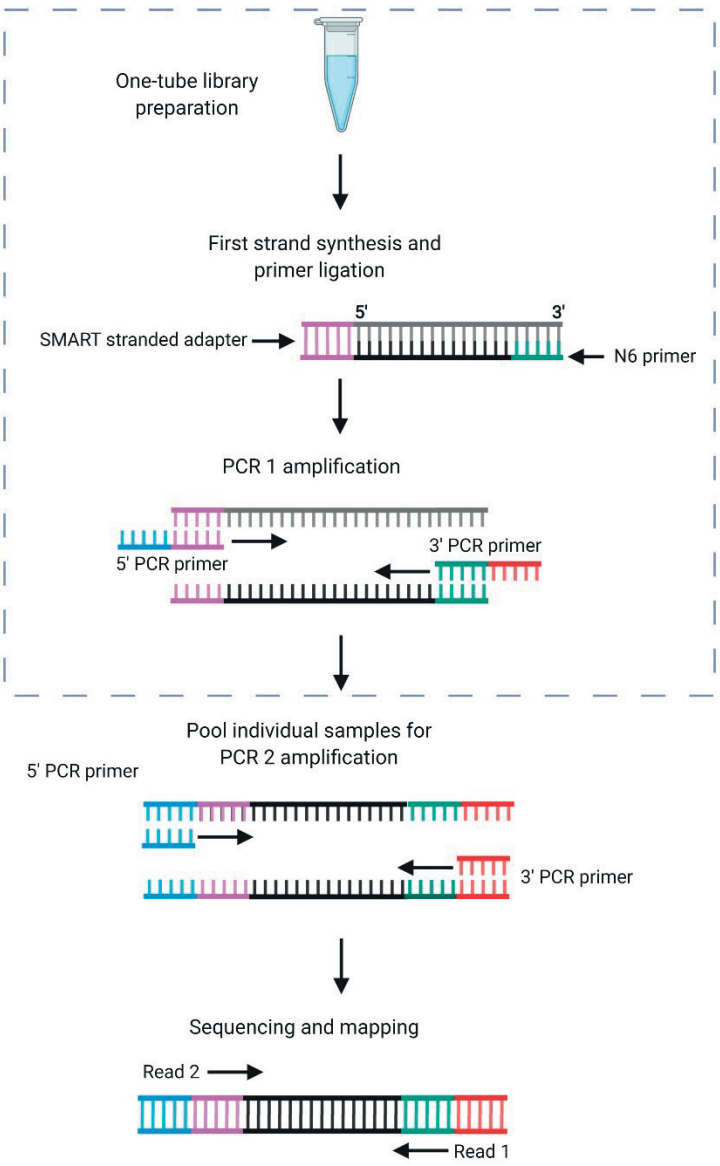


Figure 16. The SMART library preparation method was tailored to a single-tube protocol to preserve all immunoprecipitated RNA. The SMART scN6 Primer (green) is a random primer that allows the generation of cDNA from all immunoprecipitated RNA fragments. After reverse transcription, the SMART Stranded Adapter (pink) is ligated to the cDNA product. Next, full-length Illumina adapters including barcodes are added via a first round of PCR amplification (PCR 1). To minimize the number of samples to be processed downstream, individual samples can be pooled for the second round of PCR amplification (PCR 2) using primers universal to all libraries. The final library is compatible with any Illumina sequencing platform.

Insights into m6A modification in oocytes and preimplantation embryos

m6A modification in oocytes and preimplantation embryos

m6A is the most abundant internal mRNA modification and has been proven to perform multiple functions in mRNA processing, translation, and degradation (Zhao et al. 2017a; Zaccara et al. 2019). m6A is also present in noncoding RNAs (Meyer et al. 2012; Xiao et al. 2019) and carRNAs, including promoter-associated RNAs, enhancer RNAs, and retrotransposon-derived RNAs (Liu et al. 2020a; Chelmicki et al. 2021; Liu et al. 2021; Xu et al. 2021). It is deposited by m6A 'writer' methyltransferases (the METTL3 and METTL14 complex) and other associated proteins (WTAP, VIRMA, ZC3H13, CBLL1, RBM15 and RBM15B) (Roundtree et al. 2017b; Du et al. 2019). The mouse genome encodes five YTH domain-containing proteins, YTHDF1–YTHDF3, YTHDC1 and YTHDC2, that selectively bind to m6A-modified RNAs and impact their metabolism (Dominissini et al. 2012; Batista et al. 2014; Aguilo et al. 2015; Li et al. 2016; Bailey et al. 2017; Roundtree et al. 2017b; Wojtas et al. 2017; Boccaletto et al. 2018; Jain et al. 2018; Chen et al. 2021a). ALKBH5 and FTO are two identified m6A 'erasers' that remove m6A (Jia et al. 2011; Zheng et al. 2013a).

The maternal transcriptome is transcribed and accumulates during follicle growth in oogenesis and is completed in GV oocytes (Eppig and Schroeder 1989). Later, under the influence of hormonal triggers, mRNAs are gradually degraded when GV oocytes develop into MII oocytes, which are fully mature oocytes awaiting fertilization by sperm (Bachvarova et al. 1985; Paynton et al. 1988). However, it is still not fully understood why oocytes undergo gradual maternal mRNA clearance until the 2-cell stage in mice. Over the past few years, it has become clear that m6A modification plays critical roles in the degradation of maternal mRNA species (Sha et al. 2019; Vastenhouw et al. 2019).

Mettl3-knockout leads to early embryonic lethality (Geula et al. 2015a). Female mice with conditional *Mettl3* knockout showed an abnormal ovarian morphology and were sterile; the flushed oocytes were stalled at the GV stage and did not reach the two-cell stage upon fertilization attempts (Lasman et al. 2020c). In another study, METTL3 knockdown by microinjection of specific small interfering RNAs or morpholinos into GV oocytes also severely inhibited oocyte maturation by decreasing the mRNA translation efficiency, which led to defects in MZT in embryos (Sui et al. 2020).

YTHDF2 is an essential determinant of mammalian oocyte competence and early zygotic development, and loss of YTHDF2 results in failure of proper oocyte maturation (Ivanova et al. 2017). Single KO of *Ythdf1* or *Ythdf2* causes learning and memory defects as well as impaired neural development (Li et al. 2018; Shi et al. 2018b). Both YTHDC1 and YTHDC2 have an essential role in oogenesis, and their knockout during oogenesis leads to a severe hypofertility phenotype (Xiao et al. 2016; Bailey et al. 2017; Roundtree et al. 2017b; Wojtas et al. 2017; Jain et al. 2018; Kasowitz et al. 2018; Liu et al. 2021).

In **Paper III**, the IF staining results showed that m6A was mostly localized in the cytoplasm. The IF staining results suggested that high levels of m6A are present in oocytes and zygotes, that m6A signals are rapidly reduced from the zygote stage to the 2-cell embryo stage, and that m6A gradually increases from the 2-cell stage to the blastocyst stage. We also showed stage-specific m6A modulation in each transitional stage, revealing a stepwise requirement of the m6A methylome that is involved in gene regulation. These results suggest that m6A is a permissive RNA modification in oocytes and preimplantation embryos that could play important roles in early embryo development.

In **Paper I**, by profiling m6A in mouse MII oocytes and blastocysts with our scMeRIP-seq method developed in **Paper II**, we found that both metagene profiles showed a typical transcriptome-wide distribution pattern of m6A and that m6A peaks were strongly enriched in the vicinity of stop codons. The consensus motif RRACH was identified in both MII oocytes and blastocysts. Gene ontology (GO) analysis of cell-specific m6A genes revealed the expected biological relevance in oocytes and blastocysts. In MII oocytes, methylated genes were enriched in categories such as transcription, DNA templating, mRNA processing, RNA splicing and the cell cycle, suggesting that m6A is an essential modification involved in oocyte processes. In blastocysts, the enriched GO terms were categorized into transcription, DNA templating, the cell cycle, covalent chromatin modification, stem cell population maintenance, in utero embryonic development, etc., indicating that m6A-modified genes also play important roles in blastocysts. These findings suggest that m6A is a permissive modification in mouse MII oocytes and blastocysts and plays vital roles during oogenesis and embryogenesis.

m6A modification meets transposable elements

Recently, the focus of m6A studies has shifted from mRNAs to noncoding RNAs, which are termed carRNAs and include promoter-associated RNAs, enhancer RNAs and transposable element (TE)-derived RNAs (repeat RNAs) (Li et al. 2020; Liu et al. 2020a; Chen et al. 2021b;

Liu et al. 2021; Xiong et al. 2021; Xu et al. 2021). Several studies have shown that m6A is involved in the chromatin environment and transcription.

TEs constitute a substantial proportion of the mammalian genome and modulate genome structure and expression; they are also crucial for genome integrity (Goodier and Kazazian 2008). TEs are classified into retrotransposons and DNA transposons by their transposition intermediate. Retrotransposons include long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs) and long terminal repeat elements (LTRs) (Wicker et al. 2007; Chuong et al. 2017). Murine LTRs are further classified into three categories. Endogenous retroviruses (ERVs) are the third class and are the predominant superfamily among LTRs; ERVs include mouse endogenous retrovirus type-L (MuERV-L) and nonautonomous mammalian apparent LTR retrotransposons (MaLRs, sometimes referred to as the ERVL-MaLR group). Rodent MaLRs include the ancestral mammalian MLT family and rodent-specific ORR1 and MT families (Smit 1993; Hubley et al. 2016). A previous study suggested that MTA is the youngest ERVL subfamily and makes large contributions to poly(A) transcriptomes at specific stages of the germline cycle, particularly during oogenesis (Franke et al. 2017).

A recent study found that YTHDC1 facilitates the decay of a subset of these m6A-modified carRNAs, including members of the LINE family (Liu et al. 2020a). Another study showed that m6A modification on RNA exerts a protective effect on the maintenance of cellular integrity by clearing reactive ERV-derived RNA species, especially intracisternal A-particles (IAPs), a highly active subtype belonging to the ERVK family (Chelmicki et al. 2021). Notably, three recent studies reported that m6A modification is critical for maintaining cellular integrity by silencing endogenous retrovirus (ERV)-derived RNAs (Liu et al. 2021; Xiong et al. 2021; Xu et al. 2021).

In **Paper I**, we found that two retrotransposon subfamilies (L1Md_A and L1Md_T of the LINE-1 family) were enriched with m6A in mouse tissues. However, m6A was predominantly localized to the MTA_Mm subfamily, the evolutionarily youngest component of the MaLR family, in mouse MII oocytes. By comparison, m6A methylation of these three subfamilies was not found in blastocysts. A motif search showed that RRACH motifs were located throughout all the consensus sequences, suggesting the extensive m6A modification in MTA-derived transcripts. In **Paper III**, we have demonstrated that m6A is widespread in the transcriptomes of mouse oocytes and early embryos, showing stage specific marking of transcripts, and is dynamically regulated in response to developmental events. Specifically, the pronounced m6A

deposition on transcripts involved in transcriptional regulation, suggests a potential regulatory layer of m6A at the transcription level. Overall, these studies have started to reveal the roles of m6A-modified repeat elements in early embryogenesis and are expected to provide valuable insights and a deeper understanding of m6A.

Future perspectives

After the first publications of transcriptome-wide m6A mapping methods in 2012, m6A mapping methods stimulated a trend of explosive growth in research on m6A. To date, more than 2000 m6A mapping studies have been published in various contexts (Sayers et al. 2021). However, these methods require large amounts of materials, which hinders their application in many settings. In this thesis, we developed scMeRIP-seq, which allowed transcriptome-wide detection of m6A from picogram-level samples and even single cells.

We can speculate that after the development of scMeRIP-seq, it will be applied in many contexts where materials are lacking, such as research on early embryonic development and cancer. Extensive studies have shown that m6A affects the fate of modified RNA molecules and plays critical roles in almost all processes of cancer development (Huang et al. 2020). However, the contribution of m6A features to tumor heterogeneity and development remains unknown. Our methods pave the way for future application of a high-throughput droplet microfluidics platform combined with scMeRIP-seq to profile the m6A landscape of thousands of cells at single-cell resolution, an approach that could identify key differences in m6A features between normal cells and cancer cells (Baslan and Hicks 2017; Grosselin et al. 2019).

We speculate that future improvements will enhance the utility of scMeRIP-seq to analyze single somatic cells with high throughput by combining this method with a droplet microfluidics platform, optimizing fragmentation, applying automated liquid handling robots, and refining methods of data analysis. All these improvements are expected to provide superior detection accuracy, precision, and reproducibility without introducing handling bias. This method of single-cell m6A immunoprecipitation followed by sequencing will pave the way for not only uncovering the role of m6A in single cells but also revealing heterogeneity in m6A during cellular differentiation and development. In addition, we anticipate that scMeRIP-seq may be combined with existing approaches, such as single-cell analyses of genome sequences, chromatin accessibility, DNA methylation, histone modifications and chromosome conformation. All these possibilities provide opportunities for the technologies of single-cell multiomics analyses and offer approaches for integrative analysis of single cells with

multiomics data. As m6A is incorporated cotranscriptionally into RNA transcripts by the METTL3–METTL14 core methyltransferase complex, m6A on carRNAs has been found to serve as a switch to control the chromatin state and is directly involved in regulating the transcription of thousands of genes (Billon and Cristofari 2021; He and Lan 2021; Luense and Berger 2021; Wei and He 2021). scMeRIP-seq combined with single-cell genomic and epigenomic profiling technologies are anticipated to provide new mechanistic insights into the cotranscriptional interplay between m6A and other genomic factors.

In addition to studying m6A in preimplantation embryos and cancer cells, it is also possible to analyze m6A dynamics to monitor the progression of other human diseases in rare or precious biological samples during physiological or pathological processes at the single-cell level without disturbing or consuming the entire sample.

Regarding other improvements to scMeRIP-seq, it could be possible to detect m6A across the transcriptome at single-nucleotide resolution. The scMeRIP-seq protocol involves immunoprecipitation of 100∼500 nt-long RNA fragments with m6A-specific antibodies. Via this method, m6A peaks can be generated and located, but specific m6A residues cannot be identified at single-nucleotide resolution. We are also aiming to optimize the current library preparation methods of scMeRIP-seq to develop another advanced ligation-free method. In this optimized method, Illumina adapters are attached to both ends of the immunoprecipitated m6A-containing fragments, and the input size requirements of the fragments are flexible, allowing the fragments to contain fewer than approximately 50 nt. As noted in a previous study, m6A occurs in a consensus motif that is necessary for methylation, and after mapping to the reference genome, m6A-modified adenosines can be identified at single-nucleotide resolution based on the RRACH consensus motif (Schwartz et al. 2013).

At the end of the discussion, I want to suggest a bold outlook for scMeRIP-seq, predicting that it will be widely used in all single-cell sequencing research. During the past decade, single-cell RNA-seq technology has already provided fascinating insights into the processes of various developmental, physiological and disease systems. However, the findings are still not enough to solve some problems and explain some phenomena in biological studies. Increasing evidence suggests that m6A RNA modification is at the heart of gene regulation and is associated with the chromatin status. Moreover, the modifications that occur on RNA precisely control protein translation and thus regulate cell metabolism. Studying RNA modifications at the single-cell level will enable an even more profound understanding of these unresolved issues, leading to

the development of m6A atlases describing the gene transcription and expression profiles and contributing to the field of omics in biology.

References

Abakir A, Giles TC, Cristini A, Foster JM, Dai N, Starczak M, Rubio-Roldan A, Li M, Eleftheriou M, Crutchley J et al. 2020. N(6)-methyladenosine regulates the stability of RNA:DNA hybrids in human cells. *Nat Genet* **52**: 48-55.

Adams JM, Cory S. 1975. Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature* **255**: 28-33.

Aguilo F, Zhang F, Sancho A, Fidalgo M, Di Cecilia S, Vashisht A, Lee DF, Chen CH, Rengasamy M, Andino B et al. 2015. Coordination of m(6)A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. *Cell Stem Cell* **17**: 689-704.

Akhtar J, Renaud Y, Albrecht S, Ghavi-Helm Y, Roignant J-Y, Silies M, Junion G. 2020. m⁶A RNA methylation regulates promoter proximal pausing of RNA Polymerase II. *bioRxiv* doi:10.1101/2020.03.05.978163: 2020.2003.2005.978163.

Amort T, Rieder D, Wille A, Khokhlova-Cubberley D, Riml C, Trixl L, Jia XY, Micura R, Lusser A. 2017. Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol* **18**: 1.

Bachvarova R, De Leon V, Johnson A, Kaplan G, Paynton BV. 1985. Changes in total RNA, polyadenylated RNA, and actin mRNA during meiotic maturation of mouse oocytes. *Dev Biol* **108**: 325-331.

Bailey AS, Batista PJ, Gold RS, Chen YG, de Rooij DG, Chang HY, Fuller MT. 2017. The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. *Elife* **6**.

Barbieri I, Kouzarides T. 2020. Role of RNA modifications in cancer. *Nat Rev Cancer* **20**: 303-322.

Baslan T, Hicks J. 2017. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer* **17**: 557-569.

Bass BL. 2002. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**: 817-846.

Bass BL, Weintraub H. 1988. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**: 1089-1098.

Batista PJ, Molinie B, Wang J, Qu K, Zhang J, Li L, Bouley DM, Lujan E, Haddad B, Daneshvar K et al. 2014. m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**: 707-719.

Begik O, Lucas MC, Liu H, Ramirez JM, Mattick JS, Novoa EM. 2020. Integrative analyses of the RNA modification machinery reveal tissue- and cancer-specific signatures. *Genome Biol* **21**: 97.

Billon V, Cristofari G. 2021. Nascent RNA m(6)A modification at the heart of the gene-retrotransposon conflict. *Cell Res* doi:10.1038/s41422-021-00518-5.

Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, de Crecy-Lagard V, Ross R, Limbach PA, Kotter A et al. 2018. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46**: D303-D307.

Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. 1997. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* **3**: 1233-1247.

Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**: 143-146.

Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP. 2018. Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution* **9**: 410-419.

Chelmicki T, Roger E, Teissandier A, Dura M, Bonneville L, Rucli S, Dossin F, Fouassier C, Lameiras S, Bourc'his D. 2021. m(6)A RNA methylation regulates the fate of endogenous retroviruses. *Nature* **591**: 312-316.

Chen C, Liu W, Guo J, Liu Y, Liu X, Liu J, Dou X, Le R, Huang Y, Li C et al. 2021a. Nuclear m(6)A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. *Protein Cell* doi:10.1007/s13238-021-00837-8.

Chen C, Liu W, Guo J, Liu Y, Liu X, Liu J, Dou X, Le R, Huang Y, Li C et al. 2021b. Nuclear m(6)A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. *Protein Cell* **12**: 455-474.

Chen K, Lu Z, Wang X, Fu Y, Luo GZ, Liu N, Han D, Dominissini D, Dai Q, Pan T et al. 2015. High-resolution N(6) -methyladenosine (m(6) A) map using photo-crosslinking-assisted m(6) A sequencing. *Angew Chem Int Ed Engl* **54**: 1587-1590.

Chen Q, Yan W, Duan E. 2016. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat Rev Genet* **17**: 733-743.

Cho S, Lee G, Pickering BF, Jang C, Park JH, He L, Mathur L, Kim SS, Jung S, Tang HW et al. 2021. mTORC1 promotes cell growth via m(6)A-dependent mRNA degradation. *Mol Cell* **81**: 2064-2075 e2068.

Chu JM, Ye TT, Ma CJ, Lan MD, Liu T, Yuan BF, Feng YQ. 2018. Existence of Internal N7-Methylguanosine Modification in mRNA Determined by Differential Enzyme Treatment Coupled with Mass Spectrometry Analysis. *ACS Chem Biol* **13**: 3243-3250.

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* **18**: 71-86.

Clancy MJ, Shambaugh ME, Timpte CS, Bokar JA. 2002. Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. *Nucleic Acids Res* **30**: 4509-4518.

Collombet S, Ranisavljevic N, Nagano T, Varnai C, Shisode T, Leung W, Piolot T, Galupa R, Borensztein M, Servant N et al. 2020. Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**: 142-146.

Cowling VH. 2009. Regulation of mRNA cap methylation. *Biochem J* **425**: 295-302.

Crick F. 1970. Central dogma of molecular biology. *Nature* **227**: 561-563.

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY et al. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**: 548-552.

Delatte B, Wang F, Ngoc LV, Collignon E, Bonvin E, Deplus R, Calonne E, Hassabi B, Putmans P, Awe S et al. 2016. RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* **351**: 282-285.

Desrosiers R, Friderici K, Rottman F. 1974. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A* **71**: 3971-3975.

Dhote V, Sweeney TR, Kim N, Hellen CU, Pestova TV. 2012. Roles of individual domains in the function of DHX29, an essential factor required for translation of structured mammalian mRNAs. *Proc Natl Acad Sci U S A* **109**: E3150-3159.

Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M et al. 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**: 201-206.

Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WC et al. 2016. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* **530**: 441-446.

Du X, Yuan L, Wu M, Men M, He R, Wang L, Wu S, Xiang Y, Qu X, Liu H et al. 2019. Variable DNA methylation of aging-related genes is associated with male COPD. *Respir Res* **20**: 243.

Duda KJ, Ching RW, Jerabek L, Shukeir N, Erikson G, Engist B, Onishi-Seebacher M, Perrera V, Richter F, Mittler G et al. 2021. m6A RNA methylation of major satellite repeat transcripts facilitates chromatin association and RNA:DNA hybrid formation in mouse heterochromatin. *Nucleic Acids Res* doi:10.1093/nar/gkab364.

Eisenberg E, Levanon EY. 2018. A-to-I RNA editing - immune protector and transcriptome diversifier. *Nat Rev Genet* **19**: 473-490.

Enroth C, Poulsen LD, Iversen S, Kirpekar F, Albrechtsen A, Vinther J. 2019. Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. *Nucleic Acids Res* **47**: e126.

Eppig JJ, Schroeder AC. 1989. Capacity of mouse oocytes from preantral follicles to undergo embryogenesis and development to live young after growth, maturation, and fertilization in vitro. *Biol Reprod* **41**: 268-276.

Franke V, Ganesh S, Karlic R, Malik R, Pasulka J, Horvat F, Kuzman M, Fulka H, Cernohorska M, Urbanova J et al. 2017. Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res* **27**: 1384-1394.

Frye M, Harada BT, Behm M, He C. 2018. RNA modifications modulate gene expression during development. *Science* **361**: 1346-1349.

Frye M, Jaffrey SR, Pan T, Rechavi G, Suzuki T. 2016. RNA modifications: what have we learned and where are we headed? *Nature Reviews Genetics* **17**: 365-372.

Furuichi Y, Morgan M, Shatkin AJ, Jelinek W, Salditt-Georgieff M, Darnell JE. 1975. Methylated, blocked 5 termini in HeLa cell mRNA. *Proc Natl Acad Sci U S A* **72**: 1904-1908.

Gao L, Wu K, Liu Z, Yao X, Yuan S, Tao W, Yi L, Yu G, Hou Z, Fan D et al. 2018. Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution. *Cell* **173**: 248-259 e215.

Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A et al. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* **15**: 201-206.

Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, Winkler R, Nir R, Lasman L, Brandis A et al. 2019. Deciphering the "m(6)A Code" via Antibody-Independent Quantitative Profiling. *Cell* **178**: 731-747 e716.

Geula S, Moshitch-Moshkovitz S, Dominissini D, Mansour AA, Kol N, Salmon-Divon M, Hershkovitz V, Peer E, Mor N, Manor YS et al. 2015a. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **347**: 1002-1006.

Geula S, Moshitch-Moshkovitz S, Dominissini D, Mansour AA, Kol N, Salmon-Divon M, Hershkovitz V, Peer E, Mor N, Manor YS et al. 2015b. Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. *Science* **347**: 1002-1006.

Goodier JL, Kazazian HH, Jr. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* **135**: 23-35.

Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, Moritz L, Sultan C, Gurczynski SJ, Moore BB et al. 2018. A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev Cell* **46**: 651-667 e610.

Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, Dahmani A, Lameiras S, Reyal F, Frenoy O et al. 2019. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat Genet* **51**: 1060-1066.

Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J et al. 2014. The DNA methylation landscape of human early embryos. *Nature* **511**: 606-610.

Guo J, Tang HW, Li J, Perrimon N, Yan D. 2018. Xio is a component of the Drosophila sex determination pathway and RNA N(6)-methyladenosine methyltransferase complex. *Proc Natl Acad Sci U S A* **115**: 3674-3679.

Hao J, Hu H, Jiang Z, Yu X, Li C, Chen L, Xia Y, Liu D, Wang D. 2020. microRNA-670 modulates Igf2bp1 expression to regulate RNA methylation in parthenogenetic mouse embryonic development. *Sci Rep* **10**: 4782.

He C, Lan F. 2021. RNA m(6)A meets transposable elements and chromatin. *Protein Cell* doi:10.1007/s13238-021-00859-2.

Helm M, Motorin Y. 2017. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat Rev Genet* **18**: 275-291.

Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, Wu X, Wen L, Tang F, Huang Y et al. 2016. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res* **26**: 304-319.

Hsu PJ, Zhu Y, Ma H, Guo Y, Shi X, Liu Y, Qi M, Lu Z, Shi H, Wang J et al. 2017. Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis. *Cell Res* **27**: 1115-1127.

Huang H, Weng H, Chen J. 2020. m(6)A Modification in Coding and Non-coding RNAs: Roles and Therapeutic Implications in Cancer. *Cancer Cell* **37**: 270-288.

Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, Chen Z, Deng X, Xiao G, Auer F et al. 2019a. Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. *Nature* **567**: 414-419.

Huang T, Chen W, Liu J, Gu N, Zhang R. 2019b. Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nat Struct Mol Biol* **26**: 380-388.

Huang T, Guo J, Lv Y, Zheng Y, Feng T, Gao Q, Zeng W. 2019c. Meclofenamic acid represses spermatogonial proliferation through modulating m(6)A RNA modification. *J Anim Sci Biotechnol* **10**: 63.

Huber SM, van Delft P, Mendil L, Bachman M, Smollett K, Werner F, Miska EA, Balasubramanian S. 2015. Formation and abundance of 5-hydroxymethylcytosine in RNA. *Chembiochem* **16**: 752-755.

Hubley R, Finn RD, Clements J, Eddy SR, Jones TA, Bao W, Smit AF, Wheeler TJ. 2016. The Dfam database of repetitive DNA families. *Nucleic Acids Res* **44**: D81-89.

Imanishi M, Tsuji S, Suda A, Futaki S. 2017. Detection of N(6)-methyladenosine based on the methyl-sensitivity of MazF RNA endonuclease. *Chem Commun (Camb)* **53**: 12930-12933.

Ivanova I, Much C, Di Giacomo M, Azzi C, Morgan M, Moreira PN, Monahan J, Carrieri C, Enright AJ, O'Carroll D. 2017. The RNA m(6)A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence. *Mol Cell* **67**: 1059-1067 e1054.

Jain D, Puno MR, Meydan C, Lailler N, Mason CE, Lima CD, Anderson KV, Keeney S. 2018. ketu mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. *Elife* **7**.

Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang YG et al. 2011. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol* **7**: 885-887.

Kasowitz SD, Ma J, Anderson SJ, Leu NA, Xu Y, Gregory BD, Schultz RM, Wang PJ. 2018. Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet* **14**: e1007412.

Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, Haripal B, Zucker-Scharff I, Moore MJ, Park CY et al. 2015. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev* **29**: 2037-2053.

Khoddami V, Yerra A, Mosbruger TL, Fleming AM, Burrows CJ, Cairns BR. 2019. Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proc Natl Acad Sci U S A* **116**: 6784-6789.

Klemm SL, Shipony Z, Greenleaf WJ. 2019. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* **20**: 207-220.

Klungland A, Dahl JA, Greggains G, Fedorcsak P, Filipczyk A. 2016. Reversible RNA modifications in meiosis and pluripotency. *Nat Methods* **14**: 18-22.

Knuckles P, Lence T, Haussmann IU, Jacob D, Kreim N, Carl SH, Masiello I, Hares T, Villasenor R, Hess D et al. 2018. Zc3h13/Flacc is required for adenosine

methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m(6)A machinery component Wtap/Fl(2)d. *Genes Dev* **32**: 415-429.

Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**: 420-424.

Kontur C, Jeong M, Cifuentes D, Giraldez AJ. 2020. Ythdf m(6)A Readers Function Redundantly during Zebrafish Development. *Cell Rep* **33**: 108598.

Kwon J, Jo YJ, Namgoong S, Kim NH. 2019. Functional roles of hnRNPA2/B1 regulated by METTL3 in mammalian embryonic development. *Sci Rep* **9**: 8640.

Lasman L, Hanna JH, Novershtern N. 2020a. Role of m6A in Embryonic Stem Cell Differentiation and in Gametogenesis. *Epigenomes* **4**.

Lasman L, Hanna JH, Novershtern N. 2020b. Role of m(6)A in Embryonic Stem Cell Differentiation and in Gametogenesis. *Epigenomes* **4**.

Lasman L, Krupalnik V, Viukov S, Mor N, Aguilera-Castrejon A, Schneir D, Bayerl J, Mizrahi O, Peles S, Tawil S et al. 2020c. Context-dependent functional compensation between Ythdf m(6)A reader proteins. *Genes Dev* **34**: 1373-1391.

Lavi S, Shatkin AJ. 1975. Methylated simian virus 40-specific RNA from nuclei and cytoplasm of infected BSC-1 cells. *Proc Natl Acad Sci U S A* **72**: 2012-2016.

Legrand C, Tuorto F, Hartmann M, Liebers R, Jacob D, Helm M, Lyko F. 2017. Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. *Genome Res* **27**: 1589-1596.

Li M, Zhao X, Wang W, Shi H, Pan Q, Lu Z, Perez SP, Suganthan R, He C, Bjoras M et al. 2018. Ythdf2-mediated m(6)A mRNA clearance modulates neural development in mice. *Genome Biol* **19**: 69.

Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C. 2016. Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat Chem Biol* **12**: 311-316.

Li X, Xiong X, Zhang M, Wang K, Chen Y, Zhou J, Mao Y, Lv J, Yi D, Chen XW et al. 2017. Base-Resolution Mapping Reveals Distinct m(1)A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol Cell* **68**: 993-1005 e1009.

Li X, Zhu P, Ma S, Song J, Bai J, Sun F, Yi C. 2015. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat Chem Biol* **11**: 592-597.

Li Y, Xia L, Tan K, Ye X, Zuo Z, Li M, Xiao R, Wang Z, Liu X, Deng M et al. 2020. N(6)-Methyladenosine co-transcriptionally directs the demethylation of histone H3K9me2. *Nat Genet* **52**: 870-877.

Lin Z, Hsu PJ, Xing X, Fang J, Lu Z, Zou Q, Zhang KJ, Zhang X, Zhou Y, Zhang T et al. 2017. Mettl3-/Mettl14-mediated mRNA N(6)-methyladenosine modulates murine spermatogenesis. *Cell Res* **27**: 1216-1230.

Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* **12**: 767-772.

Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, Schwartz S, Mattick JS, Smith MA, Novoa EM. 2019a. Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nat Commun* **10**: 4079.

Liu HB, Muhammad T, Guo Y, Li MJ, Sha QQ, Zhang CX, Liu H, Zhao SG, Zhao H, Zhang H et al. 2019b. RNA-Binding Protein IGF2BP2/IMP2 is a Critical Maternal Activator in Early Zygotic Genome Activation. *Adv Sci (Weinh)* **6**: 1900295.

Liu J, Dou X, Chen C, Chen C, Liu C, Xu MM, Zhao S, Shen B, Gao Y, Han D et al. 2020. N (6)-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**: 580-586.

Liu J, Gao M, He J, Wu K, Lin S, Jin L, Chen Y, Liu H, Shi J, Wang X et al. 2021. The RNA m(6)A reader YTHDC1 silences retrotransposons and guards ES cell identity. *Nature* **591**: 322-326.

Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X et al. 2014. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**: 93-95.

Livneh I, Moshitch-Moshkovitz S, Amariglio N, Rechavi G, Dominissini D. 2020. The m(6)A epitranscriptome: transcriptome plasticity in brain development and function. *Nat Rev Neurosci* **21**: 36-51.

Lovejoy AF, Riordan DP, Brown PO. 2014. Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. *PLoS One* **9**: e110799.

Luense LJ, Berger SL. 2021. RNA modification to the rescue! *Cell Host Microbe* **29**: 313-315.

Matz M, Shagin D, Bogdanova E, Britanova O, Lukyanov S, Diatchenko L, Chenchik A. 1999. Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Res* **27**: 1558-1560.

Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur JJ, Chen Q et al. 2017. Reversible methylation of m(6)Am in the 5' cap controls mRNA stability. *Nature* **541**: 371-375.

Merkurjev D, Hong WT, Iida K, Oomoto I, Goldie BJ, Yamaguti H, Ohara T, Kawaguchi SY, Hirano T, Martin KC et al. 2018. Synaptic N(6)-methyladenosine (m(6)A) epitranscriptome reveals functional partitioning of localized transcripts. *Nat Neurosci* **21**: 1004-1014.

Meyer KD. 2019. DART-seq: an antibody-free method for global m(6)A detection. *Nat Methods* **16**: 1275-1280.

Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**: 1635-1646.

Miao Z, Xin N, Wei B, Hua X, Zhang G, Leng C, Zhao C, Wu D, Li J, Ge W et al. 2016. 5-hydroxymethylcytosine is detected in RNA from mouse brain tissues. *Brain Res* **1642**: 546-552.

Mikutis S, Gu M, Sendinc E, Hazemi ME, Kiely-Collins H, Aspris D, Vassiliou GS, Shi Y, Tzelepis K, Bernardes GJL. 2020a. meCLICK-Seq, a Substrate-Hijacking and RNA Degradation Strategy for the Study of RNA Methylation. *ACS Cent Sci* **6**: 2196-2208.

Mikutis S, Gu M, Sendinc E, Hazemi ME, Kiely-Collins H, Aspris D, Vassiliou GS, Shi Y, Tzelepis K, Bernardes GJL. 2020b. meCLICK-Seq, a Substrate-Hijacking and RNA

Degradation Strategy for the Study of RNA Methylation. *ACS Central Science* **6**: 2196-2208.

Molinie B, Wang J, Lim KS, Hillebrand R, Lu ZX, Van Wittenberghe N, Howard BD, Daneshvar K, Mullen AC, Dedon P et al. 2016. m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. *Nat Methods* **13**: 692-698.

Munns TW, Liszewski MK, Sims HF. 1977. Characterization of antibodies specific for N6-methyladenosine and for 7-methylguanosine. *Biochemistry* **16**: 2163-2168.

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO, Scott J. 1993. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J Biol Chem* **268**: 20709-20712.

Nishikura K. 2010. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **79**: 321-349.

Pandolfini L, Barbieri I, Bannister AJ, Hendrick A, Andrews B, Webster N, Murat P, Mach P, Brandi R, Robson SC et al. 2019. METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation. *Mol Cell* **74**: 1278-1290 e1279.

Patil DP, Chen CK, Pickering BF, Chow A, Jackson C, Guttman M, Jaffrey SR. 2016. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**: 369-373.

Paynton BV, Rempel R, Bachvarova R. 1988. Changes in state of adenylation and time course of degradation of maternal mRNAs during oocyte maturation and early embryonic development in the mouse. *Dev Biol* **129**: 304-314.

Perry RP, Kelley DE. 1974. Existence of methylated messenger RNA in mouse L cells. *Cell* **1**: 37-42.

Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, Adhikari S, Shi Y, Lv Y, Chen YS et al. 2014. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res* **24**: 177-189.

Piskol R, Peng Z, Wang J, Li JB. 2013. Lack of evidence for existence of noncanonical RNA editing. *Nat Biotechnol* **31**: 19-20.

Polson AG, Crain PF, Pomerantz SC, McCloskey JA, Bass BL. 1991. The mechanism of adenosine to inosine conversion by the double-stranded RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry* **30**: 11507-11514.

Ramanathan A, Robb GB, Chan SH. 2016. mRNA capping: biological functions and applications. *Nucleic Acids Res* **44**: 7511-7526.

Roundtree IA, Evans ME, Pan T, He C. 2017a. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **169**: 1187-1200.

Roundtree IA, Luo GZ, Zhang Z, Wang X, Zhou T, Cui Y, Sha J, Huang X, Guerrero L, Xie P et al. 2017b. YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* **6**.

Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, Erlacher M, Rossmanith W, Stern-Ginossar N, Schwartz S. 2017. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* **551**: 251-255.

Saletore Y, Meyer K, Korlach J, Vilfan ID, Jaffrey S, Mason CE. 2012. The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome Biol* **13**: 175.

Sas-Chen A, Schwartz S. 2019. Misincorporation signatures for detecting modifications in mRNA: Not as simple as it sounds. *Methods* **156**: 53-59.

Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W et al. 2021. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **49**: D10-D17.

Schaefer M, Pollex T, Hanna K, Lyko F. 2009. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res* **37**: e12.

Schumann U, Zhang HN, Sibbritt T, Pan A, Horvath A, Gross S, Clark SJ, Yang L, Preiss T. 2020. Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biol* **18**: 40.

Schwartz S. 2018. m(1)A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map. *RNA* **24**: 1427-1436.

Schwartz S, Agarwala SD, Mumbach MR, Jovanovic M, Mertins P, Shishkin A, Tabach Y, Mikkelsen TS, Satija R, Ruvkun G et al. 2013. High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**: 1409-1421.

Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, Leon-Ricardo BX, Engreitz JM, Guttman M, Satija R, Lander ES et al. 2014a. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**: 148-162.

Schwartz S, Mumbach MR, Jovanovic M, Wang T, Maciag K, Bushkin GG, Mertins P, Ter-Ovanesyan D, Habib N, Cacchiarelli D et al. 2014b. Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep* **8**: 284-296.

Sha QQ, Zhang J, Fan HY. 2019. A story of birth and death: mRNA translation and clearance at the onset of maternal-to-zygotic transition in mammalsdagger. *Biol Reprod* **101**: 579-590.

Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, Liu C, He C. 2017. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res* **27**: 315-328.

Shi H, Zhang X, Weng YL, Lu Z, Liu Y, Lu Z, Li J, Hao P, Zhang Y, Zhang F et al. 2018. m(6)A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* **563**: 249-253.

Shu X, Cao J, Cheng M, Xiang S, Gao M, Li T, Ying X, Wang F, Yue Y, Lu Z et al. 2020. A metabolic labeling method detects m(6)A transcriptome-wide at single base resolution. *Nat Chem Biol* **16**: 887-895.

Skvortsova K, Iovino N, Bogdanovic O. 2018. Functions and mechanisms of epigenetic inheritance in animals. *Nat Rev Mol Cell Biol* **19**: 774-790.

Slobodin B, Han R, Calderone V, Vrielink J, Loayza-Puch F, Elkon R, Agami R. 2017. Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell* **169**: 326-337 e312.

Smit AF. 1993. Identification of a new, abundant superfamily of mammalian LTR-transposons. *Nucleic Acids Res* **21**: 1863-1872.

Sommer B, Kohler M, Sprengel R, Seeburg PH. 1991. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **67**: 11-19.

Sommer S, Lavi U, Darnell JE, Jr. 1978. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *J Mol Biol* **124**: 487-499.

Song B, Tang Y, Chen K, Wei Z, Rong R, Lu Z, Su J, de Magalhaes JP, Rigden DJ, Meng J. 2020. m7GHub: deciphering the location, regulation and pathogenesis of internal mRNA N7-methylguanosine (m7G) sites in human. *Bioinformatics* **36**: 3528-3536.

Song J, Yi C. 2019. Reading Chemical Modifications in the Transcriptome. *J Mol Biol* doi:10.1016/j.jmb.2019.10.006.

Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, Preiss T. 2012. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* **40**: 5023-5033.

Stewart KR, Veselovska L, Kelsey G. 2016. Establishment and functions of DNA methylation in the germline. *Epigenomics* **8**: 1399-1413.

Sui X, Hu Y, Ren C, Cao Q, Zhou S, Cao Y, Li M, Shu W, Huo R. 2020. METTL3-mediated m6A is required for murine oocyte maturation and maternal-to-zygotic transition. *Cell Cycle* **19**: 391-404.

Sun H, Zhang M, Li K, Bai D, Yi C. 2019. Cap-specific, terminal N(6)-methylation by a mammalian m(6)Am methyltransferase. *Cell Res* **29**: 80-82.

Tang C, Klukovich R, Peng H, Wang Z, Yu T, Zhang Y, Zheng H, Klungland A, Yan W. 2018. ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc Natl Acad Sci U S A* **115**: E325-E333.

Vastenhouw NL, Cao WX, Lipshitz HD. 2019. The maternal-to-zygotic transition revisited. *Development* **146**.

Vilfan ID, Tsai YC, Clark TA, Wegener J, Dai Q, Yi C, Pan T, Turner SW, Korlach J. 2013. Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. *J Nanobiotechnology* **11**: 8.

Villa E, Sahu U, O'Hara BP, Ali ES, Helmin KA, Asara JM, Gao P, Singer BD, Ben-Sahra I. 2021. mTORC1 stimulates cell growth through SAM synthesis and m(6)A mRNA-dependent control of protein synthesis. *Mol Cell* **81**: 2076-2093 e2079.

Wang M, Liu X, Chang G, Chen Y, An G, Yan L, Gao S, Xu Y, Cui Y, Dong J et al. 2018. Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. *Cell Stem Cell* **23**: 599-614 e594.

Wang X, Feng J, Xue Y, Guan Z, Zhang D, Liu Z, Gong Z, Wang Q, Huang J, Tang C et al. 2017. Corrigendum: Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature* **542**: 260.

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G et al. 2014a. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**: 117-120.

Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H, He C. 2015. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**: 1388-1399.

Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. 2014b. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol* **16**: 191-198.

Wang Y, Xiao Y, Dong S, Yu Q, Jia G. 2020. Antibody-free enzyme-assisted chemical approach for detection of N(6)-methyladenosine. *Nat Chem Biol* **16**: 896-903.

Wei J, He C. 2021. Chromatin and transcriptional regulation by reversible RNA methylation. *Curr Opin Cell Biol* **70**: 109-115.

Wei J, Liu F, Lu Z, Fei Q, Ai Y, He PC, Shi H, Cui X, Su R, Klungland A et al. 2018. Differential m(6)A, m(6)Am, and m(1)A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. *Mol Cell* **71**: 973-985 e975.

Wen J, Lv R, Ma H, Shen H, He C, Wang J, Jiao F, Liu H, Yang P, Tan L et al. 2018. Zc3h13 Regulates Nuclear RNA m(6)A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Mol Cell* **69**: 1028-1038 e1026.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O et al. 2007. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* **8**: 973-982.

Wiener D, Schwartz S. 2021. The epitranscriptome beyond m(6)A. *Nat Rev Genet* **22**: 119-131.

Wojtas MN, Pandey RR, Mendel M, Homolka D, Sachidanandam R, Pillai RS. 2017. Regulation of m(6)A Transcripts by the 3'-->5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. *Mol Cell* **68**: 374-387 e312.

Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S et al. 2018. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**: 256-260.

Xia H, Zhong C, Wu X, Chen J, Tao B, Xia X, Shi M, Zhu Z, Trudeau VL, Hu W. 2018. Mettl3 Mutation Disrupts Gamete Maturation and Reduces Fertility in Zebrafish. *Genetics* **208**: 729-743.

Xiao S, Cao S, Huang Q, Xia L, Deng M, Yang M, Jia G, Liu X, Shi J, Wang W et al. 2019. The RNA N(6)-methyladenosine modification landscape of human fetal tissues. *Nat Cell Biol* **21**: 651-661.

Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, Sun HY, Li A, Ping XL, Lai WY et al. 2016. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* **61**: 507-519.

Xiong F, Wang R, Lee JH, Li S, Chen SF, Liao Z, Hasani LA, Nguyen PT, Zhu X, Krakowiak J et al. 2021. RNA m(6)A modification orchestrates a LINE-1-host interaction that facilitates retrotransposition and contributes to long gene vulnerability. *Cell Res* doi:10.1038/s41422-021-00515-8.

Xu C, Wang X, Liu K, Roundtree IA, Tempel W, Li Y, Lu Z, He C, Min J. 2014. Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nat Chem Biol* **10**: 927-929.

Xu K, Yang Y, Feng GH, Sun BF, Chen JQ, Li YF, Chen YS, Zhang XX, Wang CX, Jiang LY et al. 2017. Mettl3-mediated m(6)A regulates spermatogonial differentiation and meiosis initiation. *Cell Res* **27**: 1100-1114.

Xu W, Li J, He C, Wen J, Ma H, Rong B, Diao J, Wang L, Wang J, Wu F et al. 2021. METTL3 regulates heterochromatin in mouse embryonic stem cells. *Nature* **591**: 317-321.

Yang X, Liu QL, Xu W, Zhang YC, Yang Y, Ju LF, Chen J, Chen YS, Li K, Ren J et al. 2019. m(6)A promotes R-loop formation to facilitate transcription termination. *Cell Res* **29**: 1035-1038.

Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W et al. 2017. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res* **27**: 606-625.

Zaccara S, Ries RJ, Jaffrey SR. 2019. Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* **20**: 608-624.

Zeng Y, Wang S, Gao S, Soares F, Ahmed M, Guo H, Wang M, Hua JT, Guan J, Moran MF et al. 2018. Refined RIP-seq protocol for epitranscriptome analysis with low input materials. *PLoS Biol* **16**: e2006092.

Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q et al. 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**: 553-557.

Zhang LS, Liu C, Ma H, Dai Q, Sun HL, Luo G, Zhang Z, Zhang L, Hu L, Dong X et al. 2019a. Transcriptome-wide Mapping of Internal N(7)-Methylguanosine Methylome in Mammalian mRNA. *Mol Cell* **74**: 1304-1316 e1308.

Zhang M, Zhai Y, Zhang S, Dai X, Li Z. 2020. Roles of N6-Methyladenosine (m(6)A) in Stem Cell Fate Decisions and Early Embryonic Development in Mammals. *Front Cell Dev Biol* **8**: 782.

Zhang Z, Chen LQ, Zhao YL, Yang CG, Roundtree IA, Zhang Z, Ren J, Xie W, He C, Luo GZ. 2019b. Single-base mapping of m(6)A by an antibody-independent method. *Sci Adv* **5**: eaax0250.

Zhao BS, Roundtree IA, He C. 2017a. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**: 31-42.

Zhao BS, Wang X, Beadell AV, Lu Z, Shi H, Kuuspalu A, Ho RK, He C. 2017b. m(6)A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* **542**: 475-478.

Zhao LY, Song J, Liu Y, Song CX, Yi C. 2020. Mapping the epigenetic modifications of DNA and RNA. *Protein Cell* **11**: 792-808.

Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, Vagbo CB, Shi Y, Wang WL, Song SH et al. 2013a. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* **49**: 18-29.

Zheng G, Dahl JA, Niu Y, Fu Y, Klungland A, Yang YG, He C. 2013b. Sprouts of RNA epigenetics: the discovery of mammalian RNA demethylases. *RNA Biol* **10**: 915-918.

Zheng HX, Zhang XS, Sui N. 2020. Advances in the profiling of N(6)-methyladenosine (m(6)A) modifications. *Biotechnol Adv* **45**: 107656.

Zhong S, Li H, Bodi Z, Button J, Vespa L, Herzog M, Fray RG. 2008. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* **20**: 1278-1288.

Zhou H, Rauch S, Dai Q, Cui X, Zhang Z, Nachtergaele S, Sepich C, He C, Dickinson BC. 2019a. Evolution of a reverse transcriptase to map N(1)-methyladenosine in human messenger RNA. *Nat Methods* **16**: 1281-1288.

Zhou L, Tian S, Qin G. 2019b. RNA methylomes reveal the m(6)A-mediated regulation of DNA demethylase gene SlDML2 in tomato fruit ripening. *Genome Biol* **20**: 156.

Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, Lian Y, Fan X, Hu B, Gao Y et al. 2018. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat Genet* **50**: 12-19.

Molecular Neurobiology (2020) 57:997–1008
https://doi.org/10.1007/s12035-019-01777-5

Histone Methylations Define Neural Stem/Progenitor Cell Subtypes in the Mouse Subventricular Zone

Zhichao Zhang[1,2] · Adeel Manaf[2] · Yanjiao Li[2,3] · Sonia Peña Perez[2] · Rajikala Suganthan[2] · John Arne Dahl[2] · Magnar Bjørås[2,4] · Arne Klungland[2,3]

Received: 27 May 2019 / Accepted: 3 September 2019 / Published online: 25 October 2019

Abstract

Neural stem/progenitor cells (NSPCs) persist in the mammalian brain throughout life and can be activated in response to the physiological and pathophysiological stimuli. Epigenetic reprogramming of NPSC represents a novel strategy for enhancing the intrinsic potential of the brain to regenerate after brain injury. Therefore, defining the epigenetic features of NSPCs is important for developing epigenetic therapies for targeted reprogramming of NSPCs to rescue neurologic function after injury. In this study, we aimed at defining different subtypes of NSPCs by individual histone methylations. We found the three histone marks, histone H3 lysine 4 trimethylation (H3K4me3), histone H3 lysine 27 trimethylation (H3K27me3), and histone H3 lysine 36 trimethylation (H3K36me3), to nicely and dynamically portray individual cell types during neurodevelopment. First, we found all three marks co-stained with NSPC marker SOX2 in mouse subventricular zone. Then, CD133, Id1, Mash1, and DCX immunostaining were used to define NSPC subtypes. Type E/B, B/C, and C/A cells showed high levels of H3K27me3, H3K36me3, and H3K4me3, respectively. Our results reveal defined histone methylations of NSPC subtypes supporting that epigenetic regulation is critical for neurogenesis and for maintaining NSPCs.

Keywords Histone methylation · NSPC subtypes · Mouse subventricular zone · Neurodevelopment

Introduction

In the postnatal mammalian brain, most of the neural stem/progenitor cells (NSPCs) are spatially restricted to two specific brain regions: the subgranular zone (SGZ) in the dentate gyrus of the hippocampus and the subventricular zone (SVZ) of the lateral ventricles [1]. As the major site for NSPCs in the postnatal central nervous system (CNS), four major cell types of NSPCs have been identified in the SVZ niche: ependyma-like stem NSPCs (type E cells), quiescent or dormant NSPCs (qNSCs; type B cells), transient amplifying progenitors (TAPs; type C cells), and migrating neuronal precursors (neuroblasts; type A cells) [2, 3] (Fig. 1b). NSPCs in SVZ can be activated in response to physiological and pathophysiological stimuli, in which they initiate CNS repair and functional recovery [4]. Therefore, understanding the dynamic regulation of NSPC subtypes may provide new insight for developing novel treatment modalities for CNS diseases.

Histone modifications are post-translational modifications to histone proteins which include methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation. These modifications have biological roles and can be inherited and are referred to as epigenetic marks. Specific histone methylation marks at promoter regions affect transcription activities [5]. Generally, histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 36 trimethylation (H3K36me3) are associated with active promoters and gene bodies of actively transcribed genes. This results in increased transcription activity, whereas histone H3 lysine 27 trimethylation (H3K27me3) is linked to transcriptional

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-019-01777-5) contains supplementary material, which is available to authorized users.

✉ Magnar Bjørås
magnar.bjoras@ntnu.no

✉ Arne Klungland
arne.klungland@medisin.uio.no

[1] Institute of Neurobiology, Xi'an Jiaotong University Health Science Center, Xi'an, China

[2] Department of Microbiology, Oslo University Hospital, Rikshospitalet, Sognsvannsveien 9, 0372 Oslo, Norway

[3] Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

[4] Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Høgskoleringen 1, 7491 Trondheim, Norway

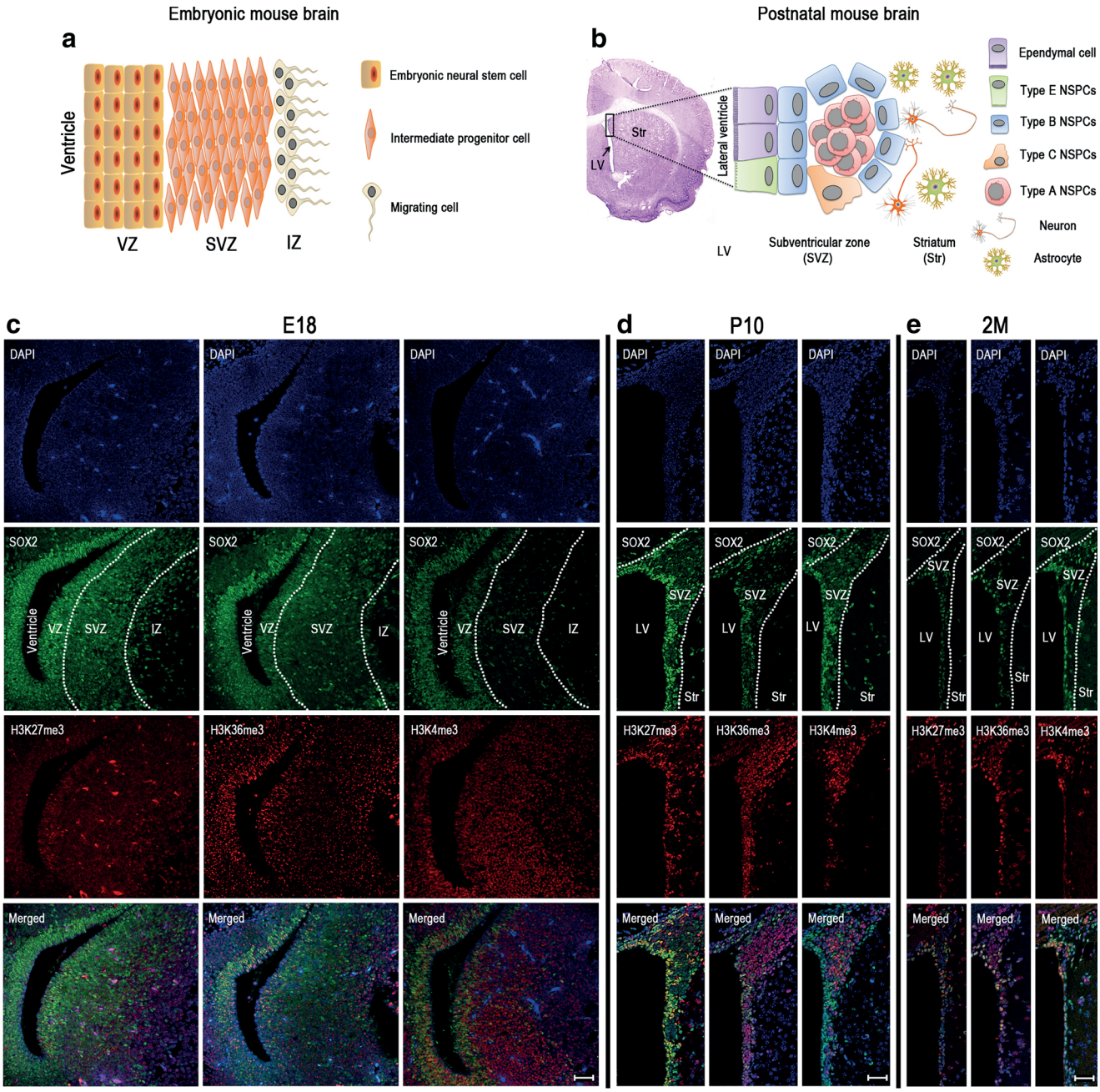


Fig. 1 H3K27me3, H3K36me3, and H3K4me3 co-located with SOX2 during neurodevelopment in SVZ. Schematics of the cell layers and cell types in the embryonic (**a**) and adult (**b**) brain. Immunofluorescent staining showed that high level of H3K27me3, H3K36me3, and H3K4me3 co-stained with SOX2 at E18 (**c**), P10 (**d**), and 2M (**e**). Nuclei were counterstained with DAPI. E18, embryo at day 18; P10, postnatal at day 10; 2M, adults 2 months. Scale bar = 50 μm

repression [6]. H3K4me3, H3K36me3, or H3K27me3 has pivotal and distinct roles in different stages of neurodevelopment and aberrant regulation of histone methylation contributes to the pathogenesis of various CNS disorders [7]. Many embryonic stem cell (ESC) promoters combine activating H3K4me3 marks and repressive H3K27me3 marks, and these bivalent domains are important dynamically regulated targets in the expression of developmental genes [8]. H3K36me3 is markedly enriched at pericentromeric heterochromatin in ESCs and fibroblasts [9]. Even though both H3K4me3 and H3K36me3 are transcriptional activators, H3K36me3 predominates in the transcribed bodies of genes, whereas nucleosomes near the transcription start site of active genes contain H3K4me3 [10]. However, we have limited understanding regarding the function of the dynamic changes in these histone methylation marks during neurodevelopment.

In this study, we observed distinct features of histone methylation in the different subtypes of NSPCs during

neurodevelopment. Type E/B cells are marked by high levels of H3K27me3, type B/C cells showed high levels of H3K36me3, and H3K4me3 is specific for type C/A cells. These results may reveal new insight into the onset of neurodevelopment and provide an innovative epigenetic signature for discovery and characterization of key regulatory genes/regions for neurogenesis.

Material and Methods

Animals

C57BL/6N mouse strain was used for this research and all mouse experiments were approved by the Animal Research Committee and the Norwegian Food Safety Authority (NFDA), and conducted in accordance with the rules and regulations of the Federation of European Laboratory Animal Science Associations (FELASA). The staff at Komparativ Medisin (KPM) Oslo University Hospital is responsible for housing and daily maintenance. Housing and environmental enrichment is according to standards. All efforts were made to minimize animal suffering and to keep the numbers of animals used to a minimum.

Method Details

P10 and adult mice were anesthetized and transcardially perfused with normal saline followed by 4% paraformaldehyde (PFA, sc-281692, Santa Cruz Biotechnology, Dallas, TX, USA). Ten milliliter normal saline and 25 ml 4% PFA were used for P10 mice, while adult mice were infused with 25 ml normal saline and 50 ml 4% PFA. For E18 mice, pregnant E18 mice were sacrificed and the fetal brains were dissected in cold PBS, and then soaked into 4% PFA for fixation. All brains were dissected and post-fixed in 4% PFA overnight at 4 °C, followed by paraffin embedding. Four-micrometer brain tissue serial slices were coronally sectioned by microtome (HM355s, Thermo Scientific, Waltham, MA, USA) and mounted onto glass slides. These sections were used for immunostaining. The slides were deparaffinized and cleared in Clear-Rite™ 3 (6901TS, Thermo Scientific) followed by rehydration in an EtOH gradient. Then the slides were heated to 95 °C in the antigen retrieval buffer (3 g sodium citrate (25114, Sigma-Aldrich), 0.4 g citric acid (251275, Sigma-Aldrich), 1000 mL H_2O, pH 6.0) for 30 min, followed by washing with 0.01 M PBS (all washes were performed three times, 5 min each). The slides were permeabilized with 0.3% Triton X-100 (T8787,

Table 1 Antibodies used during the study

Antigen	Source and host species	Concentration	Catalog no.
Anti-SOX2	Abcam, mouse monoclonal antibody	1:200	ab79351
Anti-SOX2	Abcam, rabbit polyclonal antibody	1:200	ab97959
Anti-Ki67	Abcam, rabbit polyclonal antibody	1:200	ab15580
Anti-Ki67	Invitrogen, rat monoclonal antibody	1:1000	14-5698-80
Anti-PCNA antibody	Abcam, rabbit monoclonal antibody	1:1000	ab29
Anti-PCNA antibody	Abcam, mouse monoclonal antibody	1:200	ab92552
Anti-CD133 antibody	Millipore, rat monoclonal antibody	1:100	MAB4310
Anti-ID1 antibody	R&D Systems, goat polyclonal antibody	1:200	AF4377
Anti-MASH1 antibody	Abcam, rabbit monoclonal antibody	1:1000	ab213151
Anti-MASH1 antibody	BD Pharmingen, mouse monoclonal antibody	1:100	556604
Anti-DCX antibody	Abcam, rabbit polyclonal antibody	1:200	ab18723
Anti-DCX antibody	BD Transduction Laboratories, mouse monoclonal antibody	1:200	611706
Anti-trimethyl-Histone H3 (Lys4) (H3K4me3) antibody	Millipore, rabbit monoclonal antibody	1:500	04-745
Anti-Histone H3 trimethyl Lys36 (h3k36me3) antibody	Active Motif, mouse monoclonal antibody	1:500	61021
Anti-Histone h3k27me3 (H3K27 Trimethyl) antibody	Epigentek, rabbit Polyclonal antibody	1:150	A-4039
Anti-Mouse igg (H+L) secondary antibody, Alexa Fluor 488	Invitrogen, donkey polyclonal antibody	1:500	R37114
Anti-Mouse IgG (H+L) secondary antibody, Alexa Fluor 594	Invitrogen, donkey polyclonal antibody	1:500	R37115
Anti-Rabbit IgG (H+L) secondary antibody, Alexa Fluor 488	Invitrogen, donkey polyclonal antibody	1:500	R37118
Anti-Rabbit IgG (H+L) secondary antibody, Alexa Fluor 594	Invitrogen, donkey polyclonal antibody	1:500	R37119
Anti-Rat IgG (H+L) Highly Cross-Adsorbed secondary antibody, Alexa Fluor 594	Invitrogen, donkey polyclonal antibody	1:500	A-21209
Anti-Goat IgG (H+L) Cross-Adsorbed secondary antibody, Alexa Fluor 594	Invitrogen, donkey polyclonal antibody	1:500	A-11058

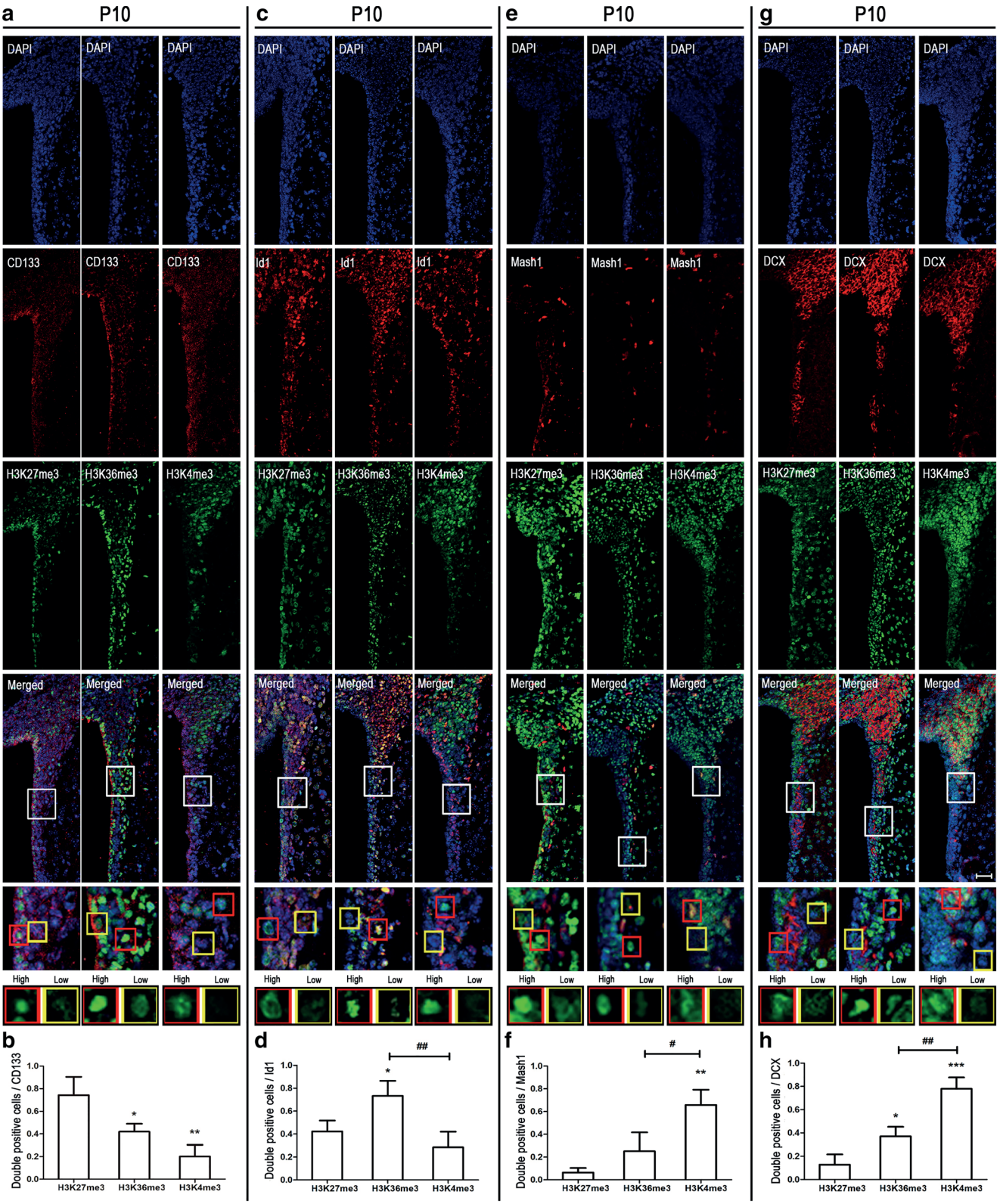


Fig. 2 Different subtypes of NSPC marker co-stained with different histone methylations at postnatal 10 days (P10). Immunocytochemical double labeling revealed different patterns of histone methylations co-stained with CD133 (**a**), Id1 (**c**), Mash1 (**e**), and DCX (**g**) respectively. Nuclei were counterstained with DAPI. Scale bar = 50 μm. The square frames are enlarged to identify a single typical high (red) and low (yellow) level cell relating to the different histone methylation mark. **b**, **d**, **f**, **h** The number of immunolabeled cells was counted for three sections in each mouse and each value represents the mean ± SD of three mice ($n = 3$). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ versus H3K27me3 group; $^{\#}P < 0.05$, $^{\#\#}P < 0.01$ versus H3K36me3 group. P10, postnatal at day 10

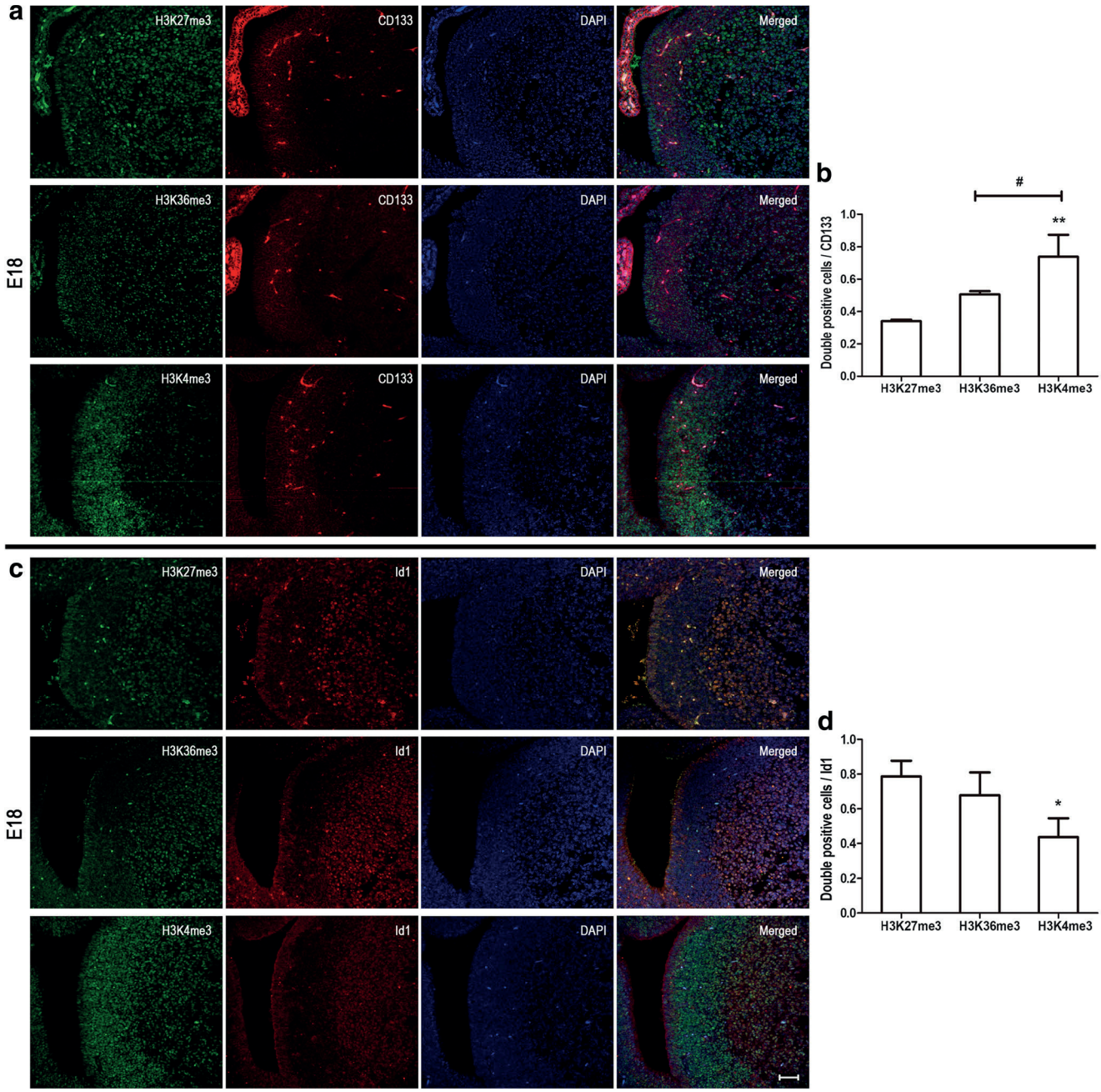


Fig. 3 CD133 and Id1 co-stained with different histone methylations at E18. Immunocytochemical double labeling revealed different patterns of histone methylations co-stained with CD133 (**a**) and Id1 (**c**), respectively. Nuclei were counterstained with DAPI. Scale bar = 50 μm. **b**, **d** The number of immunolabeled cells was counted for three sections in each mouse and each value represents the mean ± SD of three mice ($n = 3$). *$P < 0.05$, **$P < 0.01$ versus H3K27me3 group; #$P < 0.05$ versus H3K36me3 group. E18, embryo at day 18

Sigma-Aldrich, St. Louis, MO, USA) for 20 min, rinsed, and then blocked for 2 h with blocking buffer (5% normal goat serum (G9023, Sigma-Aldrich) and 5% bovine serum albumin (A7096, Sigma-Aldrich)). The samples were incubated with the primary antibodies (Table 1) overnight at 4 °C, washed with 0.01 M PBS, and then incubated with the suitable secondary antibodies. All antibodies are shown in Key resources table. The control samples were incubated in blocking buffer instead without primary antibodies. Nuclei were visualized with mounting medium including DAPI (H-1200, Vector, Burlingame, CA, USA). Images were taken with a Leica SP8 confocal microscope equipped with a × 40 oil immersion lens. The number of immunolabeled cells lining the lateral wall of the lateral ventricle was counted for three sections in each mouse, and at least three animals were used for each experiment.

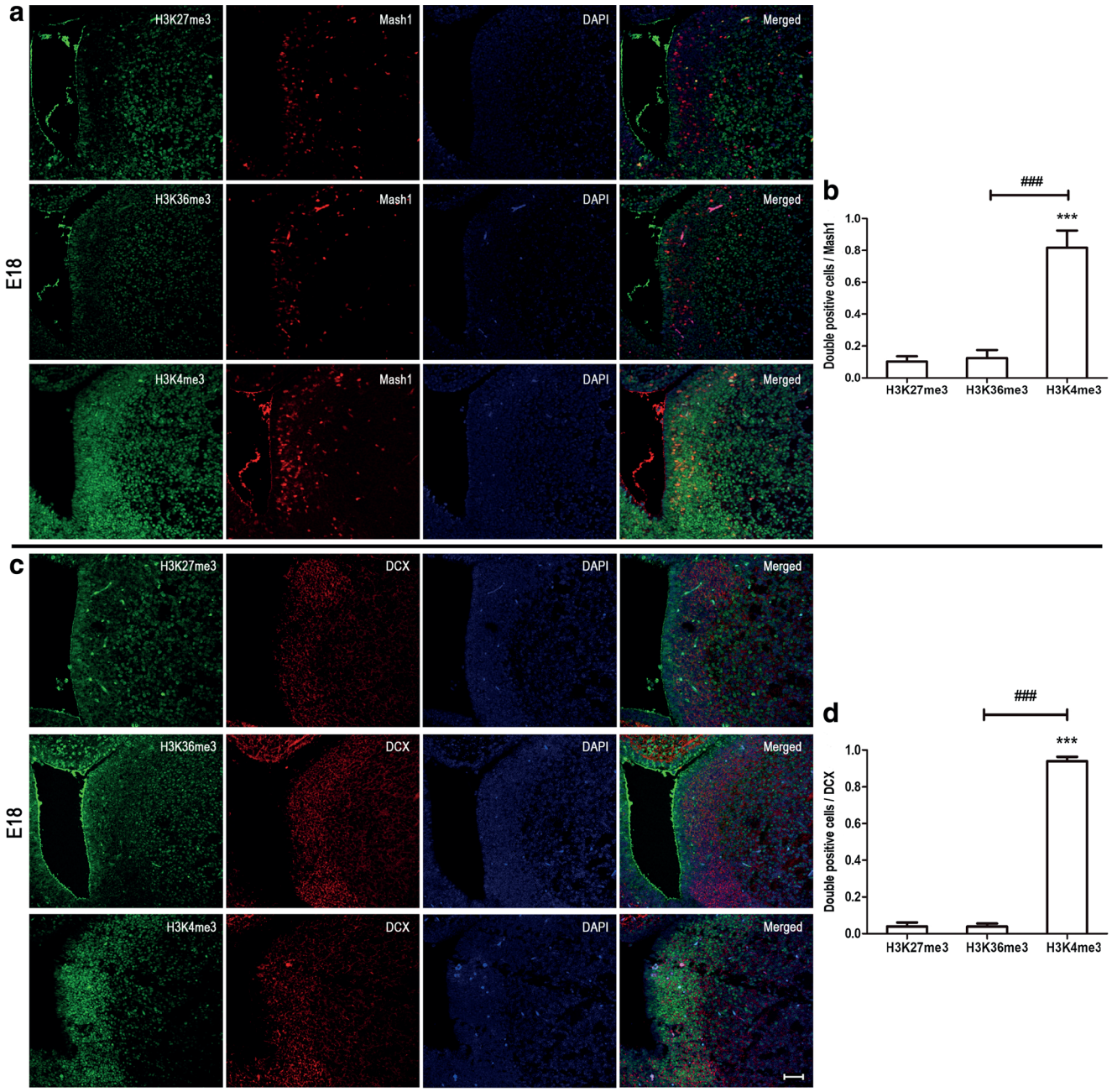


Fig. 4 Mash1 and DCX co-stained with different histone methylations at E18. Immunocytochemical double labeling showed different patterns of histone methylations in Mash1- (**a**) and DCX (**c**)-positive cells. Nuclei were counterstained with DAPI. Scale bar = 50 μm. **b**, **d** The number of immunolabeled cells was counted for three sections in each mouse and each value represents the mean ± SD of three mice ($n = 3$). ***$P < 0.001$ versus H3K27me3 group; ###$P < 0.001$ versus H3K36me3 group. E18, embryo at day 18

Quantification and Statistical Analysis

The level of histone methylation and double-positive cell was measured and defined by using Image-Pro Plus 5.1. Differences between groups were analyzed using one-way ANOVA, followed by Tukey's post hoc test. All statistical analyses were performed using GraphPad Prism 5. The data are shown as mean ± standard deviation, and $P < 0.05$ was considered as statistically significant difference.

Results

High Levels of H3K27me3, H3K36me3, and H3K4me3 in Neural Stem/Precursor Cells during Neurodevelopment in SVZ

To characterize the dynamics of histone methylations during neurodevelopment, we collected mouse brains at different time points of early life: embryo at day 18 (E18), postnatal at day 10

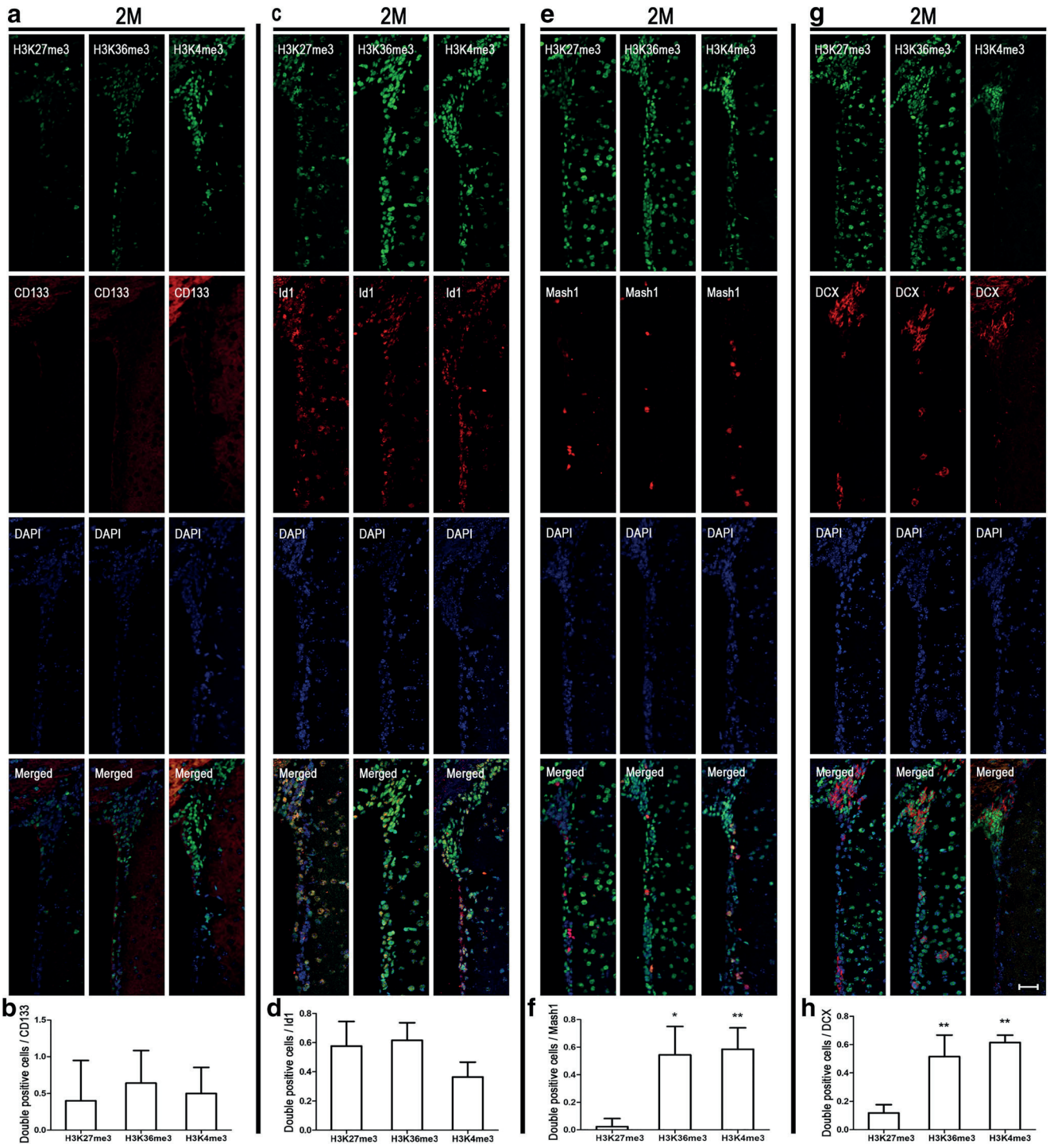
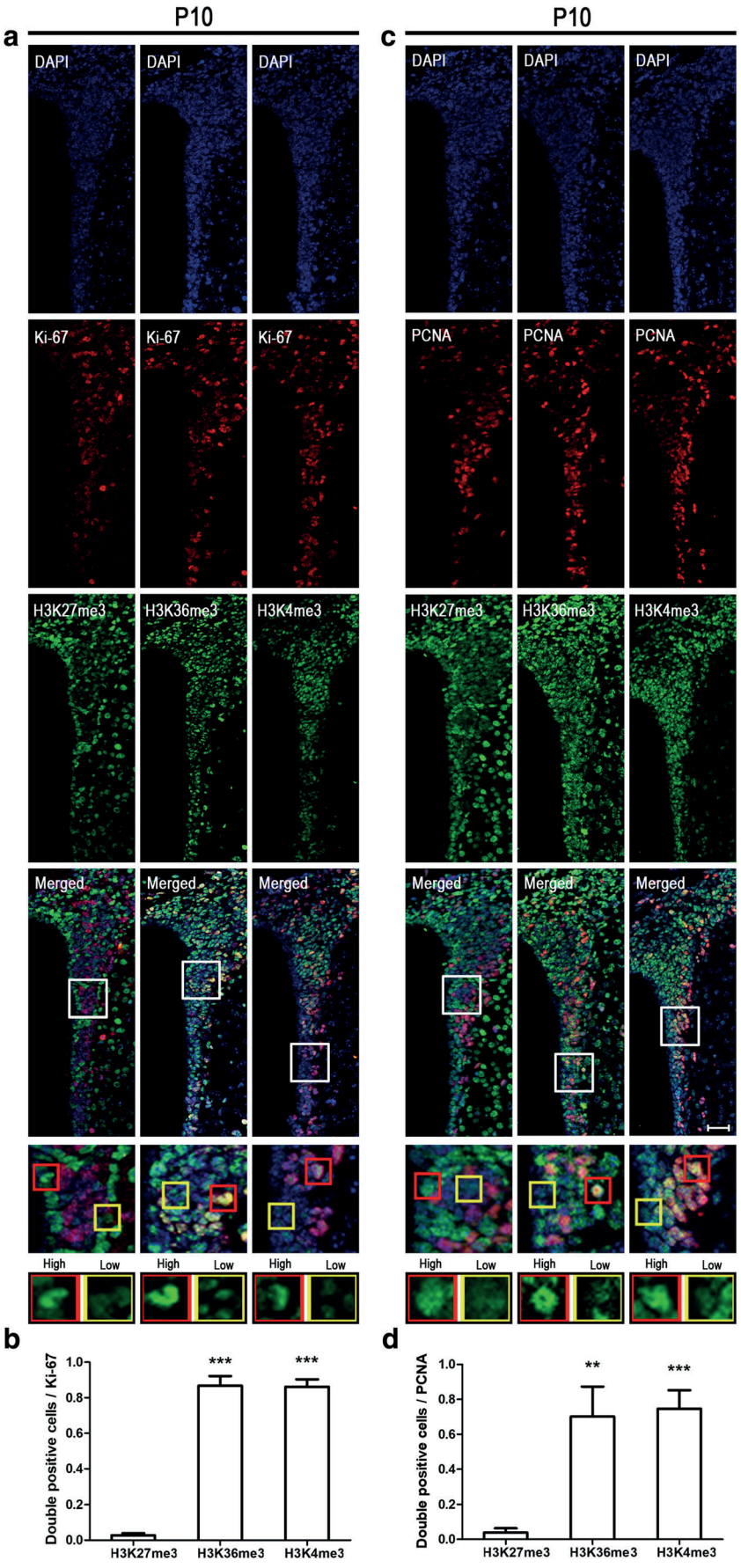


Fig. 5 Different subtypes of NSPC marker co-stained with different histone methylations at 2 months. Immunocytochemical double labeling showed different patterns of histone methylations co-stained with CD133 (**a**), Id1 (**c**), Mash1 (**e**), and DCX (**g**), respectively. Nuclei were counterstained with DAPI. Scale bar = 50 μm. **b**, **d**, **f**, and **g** The number of immunolabeled cells was counted for three sections in each mouse and each value represents the mean ± SD of three mice ($n = 3$). *$P < 0.05$, **$P < 0.01$ versus H3K27me3 group. 2M, 2 months

(P10), and adults at 2 months (2M). Then, we examined the levels of three different histone methylation marks (H3K27me3, H3K36me3, and H3K4me3) by immunofluorescence staining. All three histone marks showed strongest staining in neurogenic niches (e.g., SVZ and SGZ) during neurodevelopment, although the intensity varied among the three time points. In SVZ, the three tested histone marks showed co-localization with the established NSPC marker SOX2 at all three time points of development studied (Fig. 1c–e). In SGZ, SOX2 co-localized with H3K4me3 and H3K36me3 at all three


a
P10
DAPI
Ki-67
H3K27me3
H3K36me3
H3K4me3
Merged
High
Low
b
Double positive cells / Ki-67
1.0
0.8
0.6
0.4
0.2
0.0
***
***
H3K27me3
H3K36me3
H3K4me3
c
P10
DAPI
PCNA
H3K27me3
H3K36me3
H3K4me3
Merged
High
Low
d
Double positive cells / PCNA
1.0
0.8
0.6
0.4
0.2
0.0
**
***
H3K27me3
H3K36me3
H3K4me3

◀ **Fig. 6** High levels of H3K4me3 and H3K36me3 in cells expressing Ki-67 and PCNA at postnatal 10 days (P10). Immunocytochemical double labeling showed different patterns of histone methylations in Ki-67- (**a**) and PCNA (**c**)-positive cells. Nuclei were counterstained with DAPI. Scale bar = 50 μm. The square frames are enlarged to show the typical detail high (red) and low (yellow) levels of different histone methylation features. **b**, **d** The number of immunolabeled cells was counted for three sections in each mouse and each value represents the mean ± SD of three mice ($n = 3$). $^{**}P < 0.01$, $^{***}P < 0.001$ versus H3K27me3 group. P10, postnatal at day 10

time points, while H3K27me3 SOX2 double-positive cells were just sporadic at any time point (Supplement Fig. 1 A–D). Notably, H3K4me3, H3K36me3, and H3K27me3 stained distinct parts of SVZ, particularly at P10. H3K27me3 showed strongest staining of the ependymal cell layer, and H3K36me3 level was high in the surrounding striatal parenchyma as well as the ependymal cell layer at the lateral ventricle. In contrast, H3K4me3 staining is strongest between the ependymal cell layer and the striatal parenchyma (Fig. 1d). Previously, it has been demonstrated that in the postnatal mouse brain, type B cells locate between type A cells and the underlying striatal parenchyma as well as between type A cells and the ependymal cells, and that type C cells locate around type A cells (Fig. 1b) [11, 12]. These results suggest that histone methylation may define different subtypes of NSPCs.

High Level of H3K27me3 in CD133-Positive Cells at Early Postnatal Neurodevelopment

In postnatal mouse brain, CD133 (also known as prominin-1) is a marker for type E/B cells; Id1 marks type B/C cells (type C cells are Id1 positive, although at significantly lower levels relative to type B cells) [13]; type C cells express the highest levels of Mash1 (also known as Ascl1); and DCX marks type A cells [14]. Immunocytochemical double labeling unveils 74% CD133-positive cells showing high level of H3K27me3. On the contrary, there were few CD133-positive cells that co-stained with H3K36me3 (42%) and H3K4me3 (20%) (Fig. 2a, b). In the adult SVZ, the number of CD133-positive cell decreased markedly and there was no significant difference in immunocytochemical double labeling for the three histone methylation marks studied at this stage (Fig. 5a, b). The anatomical structure of embryo and postnatal mouse brain is noticeably different (Fig. 1a, b). In embryo mouse brain, H3K27me3 and H3K36me3 showed high levels in the ventricular zone (VZ), while high levels of H3K4me3 cells were located to SVZ (Fig. 1c). As expected, most cells in VZ and SVZ were CD133 positive. Furthermore, 74% of the CD133-positive cells co-stained with H3K4me3, while H3K27me3 and H3K36me3 showed 34% and 51% co-staining with CD133, respectively (Fig. 3a, b). Thus, it appears that high level of H3K27me3 is displayed in ependymal and quiescent neural stem cells in the SVZ (type B/E) at early postnatal neurodevelopment.

High Level of H3K36me3 in Id1-Positive Cells at Early Postnatal Neurodevelopment

Similarly, we used immunocytochemical double labeling to identify colocalization of Id1 and the three histone methylation marks. Seventy-three percent Id1-positive cells co-stained with H3K36me3, significantly higher than H3K27me3 (42%) and H3K4me3 (29%) (Fig. 2c, d). Analogous to the CD133 staining, the number of Id1-positive cells was reduced dramatically at adulthood and double labeling revealed minor differences for the three histone methylation marks (Fig. 5c, d). In embryo mouse brain, most of the Id1-positive cells were located in VZ and the majority of Id1-positive cells co-stained with H3K27me3 (79%) and H3K36me3 (68%) (Fig. 3c, d). However, just 44% H3K4me3-positive cells co-stained with Id1. These phenomena may indicate that H3K36me3 is a good marker for quiescent and active neural stem cells (type B/C) at early postnatal neurodevelopment.

High Level of H3K4me3 in Mash1 and DCX-Positive Cells at Postnatal Neurodevelopment

Mash1 (also known as Ascl1) is characterized as a proneural transcription factor and typically used as a type C cell marker. DCX is expressed in the last stage before NSPCs are migrating through the rostral migratory stream (RMS) [14]. Therefore, Mash1 and DCX were used for labeling type C and A cell, respectively. Immunocytochemical double labeling identified 66% of Mash1-positive cells co-staining with H3K4me3 at P10, while very low co-staining was observed for H3K27me3 (6%) and H3K36me3 (25%) (Fig. 2 E, F). Embryonic brain staining results showed that Mash1-positive cell appeared in SVZ; and it was similar to P10 with 82% H3K4me3, and very low H3K27me3 and H3K36me3 co-staining, 10% and 13% respectively (Fig. 4a, b). At adulthood, 58% Mash1-positive cells co-stained with H3K4me3, 2% with H3K27me3, and 54% with H3K36me3 (Fig. 5e, f). Then, double immunostaining was also used for detecting DCX and different histone methylations. During neurodevelopment, the number of H3K4me3 DCX double-positive cells was significantly higher compared with H3K27me3 or H3K36me3 double-positive cells (Fig. 2g, h; Fig. 4c, d; and Fig. 5g, h). Thus, both type C and type A are represented by H3K4me3.

H3K4me3- and H3K36me-Positive Cells Co-Stain with Proliferation Markers at Early Neurodevelopment

To further evaluate histone methylation in the proliferation state of early developmental cells in SVZ, co-staining with

the proliferation markers Ki-67 and PCNA was analyzed. We identified a noticeable difference with most of Ki-67-positive cells co-staining with H3K36me3 (87%) or H3K4me3 (86%), while only 3% of H3K27me3-positive cells co-stained with Ki-67 at P10 (Fig. 6a, b). Similarly, just 4% of PCNA-positive cells co-stained with H3K27me3 compared with 70% for H3K36me3 and 75% for H3K4me3 (Fig. 6c, d). These results strongly indicate that high levels of H3K36me3 and H3K4em3 correlate very well with proliferating cells in SVZ at early postnatal neurodevelopment.

Discussion

Traditional therapies for CNS diseases are limited. For example, treatment for clinical stroke by the administration of tissue plasminogen activator and the recent introduction of mechanical thrombectomy can only be used in a limited proportion of patients due to time constraints [15]. Accordingly, continuing efforts are in need to develop novel, safe, and more optimal and effective therapeutic strategies for CNS diseases. The dynamic regulation of histone methylations and chromatin remodeling plays essential roles in development, cellular differentiation, and cell fate maintenance [16]. More importantly, emerging evidence supports the involvement of histone methylation in the pathogenesis of CNS damage and several neurodegenerative diseases [17, 18]. In this study, we reveal how different histone methylation marks are dynamically regulated during NSPC differentiation in the mouse SVZ area, represented as marked differences in histone methylations between quiescent and active NSPCs (Fig. 7). As NSPCs can be activated by CNS damage and participate in CNS repair and functional recovery, our study may bring a novel perspective to a therapeutic strategy of CNS diseases and provide a potential histone methylation features for screening and identifying key therapeutic genes for CNS diseases.

SOX2 maintain stemness of NSPCs in a slowly proliferating stem cell state by repressing the cell cycle regulator cyclin D1 during cortex development [19]. When NSPCs enter the stage of differentiation, the levels of SOX2 decrease, which releases this repression and thus promotes cell cycle re-entry and NPC proliferation [20]. In this study, the SOX2 staining results showed that the number of SOX2 cells in SVZ gradually decreased during neurodevelopment. Notably, most cells with high level of H3K27me3 showed high level of SOX2, whereas H3K36me3 cells presented low co-staining with SOX2 cells. Furthermore, H3K4me3 and SOX2 co-staining are rare in SVZ. There is a positive correlation between the expression of SOX2 and stemness of NSPCs [21]. Thus, our results define histone methylations specific for SOX2-positive NSPCs. Moreover, we reveal that high levels of H3K27me3 exist in the early stage of NSPC development; H3K36me3 is characteristic of metaphase while H3K4me3 is enriched in the mid and later stages of NSPC development.

In the mammalian embryo brain, the proliferative region comprises two distinct zones: VZ, which is a neuroepithelial layer directly adjacent to the ventricular lumen, and SVZ, which is positioned superficial to the ventricular zone [22] (Fig. 1a). Radial glial cells (RGCs, one type of embryonic neural stem cells) reside in the VZ and generate both intermediate progenitor cells (IPCs, one type of embryonic neural precursor cells) and cortical neurons. IPCs migrate away from

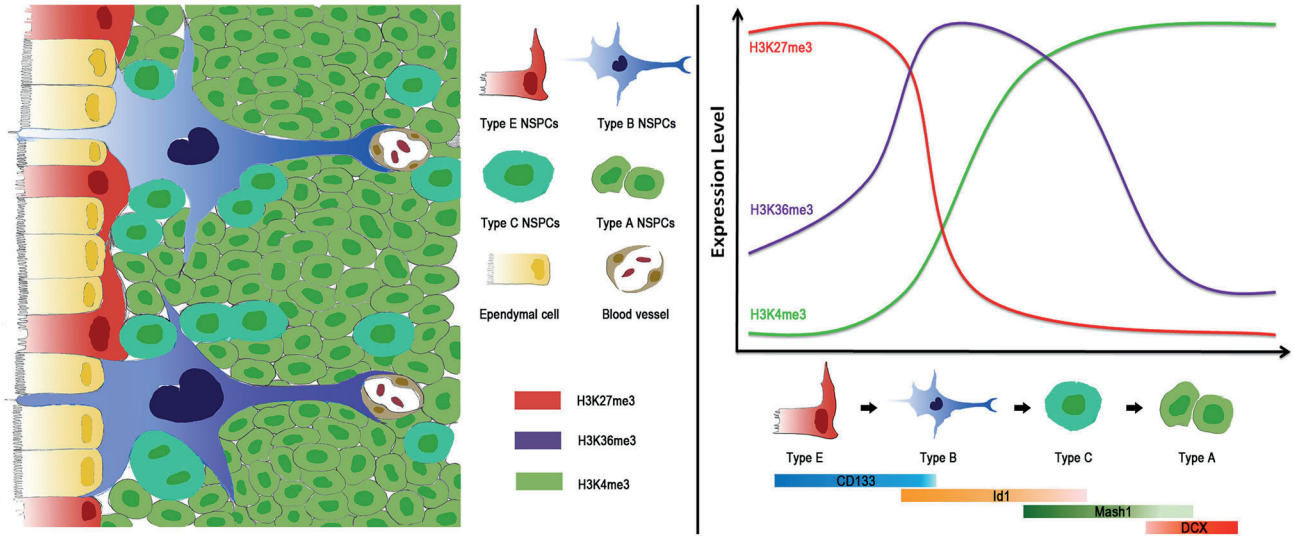


Fig. 7 Schematic model of the developmental process of NSPCs projected from this study. Histone methylations are dynamically changed during NSPC differentiation in mouse SVZ area. Different subtypes of NSPCs represented different patterns of histone methylations. Specifically, type E/B cells are marked by high levels of H3K27me3, type B/C cells showed high levels of H3K36me3, and H3K4me3 is specific for type C/A cells

the ventricular surface and establish the SVZ [23]. Therefore, the cellular composition is different in VZ and SVZ. RGCs are mostly concentrated in VZ and most of IPCs located in SVZ. In this study, embryo brain staining showed high levels of H3K27me3 and H3K36me3 in VZ, and of H3K4me3 in SVZ. Thus, it suggests high level of H3K27me3 and H3K36me3 at early stage of embryonic neural stem cell development, and of H3K4me3 at middle/late stage.

Further, our results identify significant differences among immunocytochemical double labeling in the P10 SVZ. However, we found that these distinct features disappeared in 2 months or E18; the number of NSPCs in SVZ was significantly decreased during neuronal development, and the dynamics of histone methylations described here might be one of the mechanisms underlying this regulation and might encode the difference between embryonic NSPCs and adult NSPCs. One major difference between adult and embryonic neural stem cells is their different number and their ability to differentiate into various cell types. Embryonic NSPCs can divide asymmetrically to generate neurons directly or indirectly through intermediate progenitor cells and oligodendrocytes. More importantly, at the end of the embryonic development, embryonic NSPCs begin to detach from the apical side and convert into astrocytes. Even if adult NSPCs can continue to generate neurons and oligodendrocytes, they cannot differentiate into astrocytes [24]. Histone methylation introduces epigenetic modifications with close ties to transcription and has been directly linked to lifespan regulation in many organisms [25]. For example, upon differentiation towards the neuronal lineage, some bivalent genes became expressed and lost the H3K27me3 mark, whereas those that were silenced lost the H3K4me3 and retained H3K27me3 [26]. Therefore, it is not unlikely that the embryonic and adult NSPC states are maintained by differential histone methylation profiles.

Chromatin, the template for epigenetic regulation, is a highly dynamic entity that is constantly reshaped during neurodevelopment [27]. Epigenetic regulation by histone methylation provides the necessary plasticity for cells to respond to environmental and positional cues, enabling the maintenance of acquired information without changing the DNA sequence. In this study, we showed different subtypes of NSPCs represented different features of histone methylations. These results may reveal novel insight into the onset of neurodevelopment and provide an innovative epigenetic signature for discovery and characterization of key regulatory genes for neurogenesis. However, further studies, especially whole epigenome analysis and histone profiling, are necessary for in-depth understanding of the role for individual histone methylation domains in neurodevelopment.

Acknowledgments We thank the animal facility at Oslo University Hospital for mouse handling.

Funding Information This work was funded by the Norwegian Cancer Society (to A.K. and M.B.), the Norwegian Research council, the Health Authority South East, China Postdoctoral Science Foundation (No. 2017M623152, 2018T111035, to Z.Z.) and National Natural Science Foundation of China (No. 81901156, to Z.Z.).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Ming GL, Song HJ (2011) Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron 70(4): 687–702. https://doi.org/10.1016/j.neuron.2011.05.001
2. Zhao CM, Deng W, Gage FH (2008) Mechanisms and functional implications of adult neurogenesis. Cell 132(4):645–660. https://doi.org/10.1016/j.cell.2008.01.033
3. Coskun V, Wu H, Blanchi B, Tsao S, Kim K, Zhao J, Biancotti JC, Hutnick L et al (2008) CD133+ neural stem cells in the ependyma of mammalian postnatal forebrain. Proc Natl Acad Sci U S A 105(3):1026–1031. https://doi.org/10.1073/pnas.0710000105
4. Ma DK, Bonaguidi MA, Ming GL, Song HJ (2009) Adult neural stem cells in the mammalian central nervous system. Cell Res 19(6):672–682. https://doi.org/10.1038/cr.2009.56
5. Pedersen MT, Helin K (2010) Histone demethylases in development and disease. Trends Cell Biol 20(11):662–671. https://doi.org/10.1016/j.tcb.2010.08.011
6. Tessarz P, Kouzarides T (2014) Histone core modifications regulating nucleosome structure and dynamics. Nat Rev Mol Cell Biol 15(11):703–708. https://doi.org/10.1038/nrm3890
7. Urdinguio RG, Sanchez-Mut JV, Esteller M (2009) Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. Lancet Neurol 8(11):1056–1072. https://doi.org/10.1016/S1474-4422(09)70262-5
8. Voigt P, Tee WW, Reinberg D (2013) A double take on bivalent promoters. Genes Dev 27(12):1318–1338. https://doi.org/10.1101/gad.219626.113
9. Chantalat S, Depaux A, Hery P, Barral S, Thuret JY, Dimitrov S, Gerard M (2011) Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. Genome Res 21(9):1426–1437. https://doi.org/10.1101/gr.118091.110
10. Abed JA, Jones RS (2012) H3K36me3 key to polycomb-mediated gene silencing in lineage specification. Nat Struct Mol Biol 19(12): 1214–1215. https://doi.org/10.1038/nsmb.2458
11. Ming GL, Song H (2005) Adult neurogenesis in the mammalian central nervous system. Annu Rev Neurosci 28:223–250. https://doi.org/10.1146/annurev.neuro.28.051804.101459
12. Wang YZ, Plane JM, Jiang P, Zhou CJ, Deng W (2011) Concise review: quiescent and active states of endogenous adult neural stem cells: identification and characterization. Stem Cells 29(6):907–912. https://doi.org/10.1002/stem.644

13. Nam HS, Benezra R (2009) High levels of Id1 expression define B1 type adult neural stem cells. Cell Stem Cell 5(5):515–526. https://doi.org/10.1016/j.stem.2009.08.017
14. Mamber C, Kozareva DA, Kamphuis W, Hol EM (2013) Shades of gray: the delineation of marker expression within the adult rodent subventricular zone. Prog Neurobiol 111:1–16. https://doi.org/10.1016/j.pneurobio.2013.07.003
15. Yong-Quan G, Lim YA, Sobey CG, Dheen T, Fann DYW, Arumugam TV (2018) Epigenetic regulation of inflammation in stroke. Ther Adv Neurol Diso 11:175628641877181510.1177/1756286418771815
16. Greer EL, Shi Y (2012) Histone methylation: a dynamic mark in health, disease and inheritance. Nat Rev Genet 13(5):343–357. https://doi.org/10.1038/nrg3173
17. Mehler MF (2008) Epigenetic principles and mechanisms underlying nervous system functions in health and disease. Prog Neurobiol 86(4):305–341. https://doi.org/10.1016/j.pneurobio.2008.10.001
18. Berson A, Nativio R, Berger SL, Bonini NM (2018) Epigenetic regulation in neurodegenerative diseases. Trends Neurosci 41(9):587–598. https://doi.org/10.1016/j.tins.2018.05.005
19. Hagey DW, Klum S, Kurtsdotter I, Zaouter C, Topcic D, Andersson O, Bergsland M, Muhr J (2018) SOX2 regulates common and specific stem cell features in the CNS and endoderm derived organs. PLoS Genet 14(2):e1007224. https://doi.org/10.1371/journal.pgen.1007224
20. Hagey DW, Muhr J (2014) Sox2 acts in a dose-dependent fashion to regulate proliferation of cortical progenitors. Cell Rep 9(5):1908–1920. https://doi.org/10.1016/j.celrep.2014.11.013
21. Baser A, Skabkin M, Kleber S, Dang YL, Balta GSG, Kalamakis G, Gopferich M, Ibanez DC et al (2019) Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. Nature 566(7742):100. https://doi.org/10.1038/s41586-019-0888-x
22. Noctor SC, Martinez-Cerdeno V, Kriegstein AR (2007) Contribution of intermediate progenitor cells to cortical histogenesis. Arch Neurol 64(5):639–642. https://doi.org/10.1001/archneur.64.5.639
23. Noctor SC, Martinez-Cerdeno V, Ivic L, Kriegstein AR (2004) Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nat Neurosci 7(2):136–144. https://doi.org/10.1038/nn1172
24. Kriegstein A, Alvarez-Buylla A (2009) The glial nature of embryonic and adult neural stem cells. Annu Rev Neurosci 32:149–184. https://doi.org/10.1146/annurev.neuro.051508.135600
25. Sen P, Shah PP, Nativio R, Berger SL (2016) Epigenetic mechanisms of longevity and aging. Cell 166(4):822–839. https://doi.org/10.1016/j.cell.2016.07.050
26. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2):315–326. https://doi.org/10.1016/j.cell.2006.02.041
27. Bale TL (2015) Epigenetic and transgenerational reprogramming of brain development. Nat Rev Neurosci 16(6):332–344. https://doi.org/10.1038/nrn3818