ROLES OF TET-ELEVEN TRANSLOCATION (TET) METHYLCYTOSINE DIOXYGENASES IN MAMMALIAN HEART DEVELOPMENT

A Dissertation

by

SHAOHAI FANG

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yun Huang
Committee Members,	Fen Wang
	David J. Reiner
	Yu Liu
Head of Department,	Carol Vargas

December 2021

Major Subject: Medical Sciences

ABSTRACT

DNA methylation and demethylation cycle plays critical roles in spatiotemporally orchestration of gene expression networks during mammalian development and diseases. Ten-eleven translocation (Tet) methylcytosine dioxygenases (Tet1, Tet2 and Tet3) are the major players for mediating DNA demethylation. However, the function of Tet enzymes in cardiac development is still unclear.

In this study, we took advantages of the *in vitro* Tet deficient mouse embryonic stem cell (mESCs) cardiac differentiation model and the cardiac-specific Tet deficient mice models to evaluate Tet mediated epigenetic regulation in heart development. State-of-the-art next generation sequencing (NGS) technologies in combination with cutting-edge bioinformatic analysis and sophisticated histology analysis were applied to uncover the underlying molecular mechanisms.

Our data showed that deletion of Tet proteins impaired the cardiac lineage specification during mESCs differentiation. Loss of Tet2 and Tet3 during heart development caused the non-compaction cardiomyopathy (NCC). Tet2 and Tet3 mediated chromatin accessibility enabling chromatin binding of transcription factor Ying-Yang1(YY1), and therefore facilitating enhancer-promoter interactions of genes that are important for cardiac development. In addition, cardiac specific Tet1/2/3 loss-of-function delayed the development of progenitors in the second heart field and blocked their differentiation toward cardiomyocytes, which resulted in severe cardiac development defects, including shorter outflow tract (OFT), atrial septal defect (ASD) and ventricular septal defect (ASD).

This study elaborated the crucial roles of Tet mediated DNA demethylation in heart development and provided additional molecular foundation for the prevention and treatment of congenital heart diseases in human.

DEDICATION

I would like to dedicate this dissertation to my father, my mother, and my younger sister for their selfless love.

ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. Yun Nancy Huang, for her mentorship. As the first graduate student in her lab, I really enjoyed the privilege of her close guidance. This not only made our research project succeed but also established my scientific thinking which will benefit all my independent research career.

I would like to thank my committee members, Dr. Fen Wang, Dr. David J. Reiner and Dr. Yu Liu for their guidance and support throughout the course of this research. Thank all my collaborators, including Dr. James F. Martin and his lab members, our bioinformatician Dr. Jia Li, Dr. Yubin Zhou and his lab members, and Dr. Linglin Xie and her lab members. Nothing can be achieved without their support.

Thanks also go to all the present and previous Huang lab members, all the present and previous members of the Center for Epigenetics & Disease Prevention in Institute of Biosciences & Technology (IBT), Texas A&M University, IBT Flowcytometry Core, IBT Advanced Image Core and IBT graduate program for providing substantial support for this project and my graduate study.

Finally, thanks to my family and my girlfriend for their support and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Fen Wang and Dr. David J Reiner of the Institute of Biosciences & Technology and Dr. Yu Liu of the Department of Biology & Biochemistry in University of Houston.

The NGS data analyzed for this dissertation was provided by Dr. Jia Li.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

IBT Spring Travel Award supported the student to present this study in AHA conference.

This work was also made possible in part by National Institute of Health under Grant Number R01HL134780, R01HL146852, Cancer Prevention and Research Institute of Texas under Grant Number RR140053, American Heart Association under Grant Number 16IRG27250155, and the Texas A&M University start-up funds. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the awarding offices.

NOMENCLATURE

Tet	Ten-eleven translocation
mESCs	Mouse embryonic stem cells
NGS	Next generation sequencing
NCC	Non-compaction cardiomyopathy
OFT	Outflow tract
ASD	Atrial septal defect
VSD	Ventricular septal defect
CEDP	Center for Epigenetics & Disease Prevention
IBT	Institute of Biosciences & Technology
DMNT	DNA methyltransferase
DSBH	Double-stranded β-helix
α-KG	α-ketoglutarate
5mC	5-methylcytosine
5hmC	5-hydroxymethylation
5caC	5-carboxylcytosine
5fC	5-formylcytosine
TDG	Thymine DNA glycosylase
hESCs	Human embryonic stem cells
Sfrp4	Secreted frizzled-related protein 4
PGCs	Primordial germ cells
KO	Knock-out

DKO	Tet enzymes double knock-out
Tet-TKO	Tet enzymes triple knock-out
CD	Catalytic domain
scRNA-seq	Single-cell RNA-seq
qPCR	Real-time quantitative PCR
CRISPR	Clustered regularly interspaced short palindromic repeats
cTNT	Troponin
EB	Embryonic body
DHMR	Differential hydroxymethylated regions
GREAT	Genomic Regions Enrichment of Annotations Tool
TF	Transcriptional factors
MMRRC	Mutant Mouse Resource & Research Centers
MEF	Mouse embryonic fibroblasts
LIF	Leukemia Inhibitory factor
YY1	Ying-yang1
CS	Carnegie stage
E9.5	Embryonic day 9.5
DMR	Differential methylated region
CM	Cardiomyocyte
DORV	Double outlet right ventricle
DEG	Differentially expressed genes
GO	Gene ontology

ENCODE	The encyclopedia of DNA elements
GSEA	Gene set enrichment analysis
TSS	Transcription start sites
ATAC-Seq	Assay for Transposase Accessible Chromatin with high-throughput sequencing
PET	Paired-ended tags
IACUC	Institutional Animal Care Use Committee
IHC	Immunohistochemistry
IF	Immunofluorescence
H&E	Hematoxylin-eosin
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
PBS	Phosphate buffered saline
EMT	Epithelial to mesenchymal transition
FHF	First heart field
SHF	Second heart field
WT	Wild type
CHD	Congenital heart diseases

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

CONTRIBUTORS AND FUNDING SOURCES vi

NOMENCLATURE vii

TABLE OF CONTENTS x

LIST OF FIGURES xiii

LIST OF TABLES xv

1. INTRODUCTION 1

1.1. DNA Methylation and Demethylation 1
1.2. Tet Enzymes and Active DNA Demethylation 2
1.3. Tet Enzymes Mediated DNA Demethylation in Embryonic Stem Cells 4
1.4. Tet Enzymes Mediated DNA Demethylation in Mammalian Embryonic Development 6
1.5. Tet Enzymes Mediated DNA Demethylation in Heart Development 7

2. TET ABLATION IMPAIRS MOUSE EMBRYONIC STEM CELL CARDIAC DIFFERENTIATION 10

2.1. Summary 10
2.2. Introduction 10
2.3. Results 11
2.3.1. Dynamic DNA hydroxymethylation changes during mESC-to-cardiac progenitor differentiation. 11
2.3.2. Tet Triple Knockout (TKO) mESC Exhibits Defects During Its Differentiation Toward Cardiac Progenitors. 13
2.3.3. Single-cell RNA-seq (scRNA-seq) Analysis Reveals Abnormal Cardiac Progenitor Lineage Specification in Tet-TKO mESCs. 16
2.3.4. Tet Deficiency Impairs DNA Methylation at the Hand1 Promoter. 19

2.3.5. Tet1 Is Required for Hand1 Expression During Cardiac Progenitor Differentiation..... 21
2.4. Discussions..... 24
2.5. Materials and Methods..... 26
2.5.1. Mouse Embryonic Stem Cell (mESC) Culture and Differentiation..... 26
2.5.2. Antibodies..... 26
2.5.3. Generating Tet Knockout mESCs..... 27
2.5.4. Establishing Flag-Tet1CD or dCas9-Tet1CD Expressing mESCs..... 27
2.5.5. Genomic DNA, Total RNA and Protein Isolation..... 28
2.5.6. 5hmC Dot-Blot Assay..... 28
2.5.7. Real-Time Quantitative PCR (qPCR) Assay..... 29
2.5.8. Western Blot Analysis..... 30
2.5.9. Loci Specific Methylation Analysis..... 31
2.5.10. Flow Cytometry Analysis..... 31
2.5.11. Single Cell RNA-seq Library Preparation and Data Analysis..... 32
2.5.12. CMS-IP Library Construction and Data Analysis..... 33
2.5.13. Accession Numbers and Data Availability..... 34

3. TET INACTIVATION DISRUPTS YY1 BINDING AND LONG-RANGE CHROMATIN INTERACTIONS DURING EMBRYONIC HEART DEVELOPMENT[39]..... 35

3.1. Summary..... 35
3.2. Introduction..... 35
3.3. Results..... 37
3.3.1. 5mC and 5hmC Undergo Dynamic Changes During Embryonic Heart Development in Both Human and Mice..... 37
3.3.2. Cardiac-Specific Tet Depletion Led to Ventricular Non-Compaction Cardiomyopathy (NCC) in Mice..... 40
3.3.3. Tet2 and Tet3 Regulates Cardiac-Specific Transcription During Heart Development..... 43
3.3.4. Tet2/3 Depletion Causes a Global Decrease of 5hmC with Global 5mC Levels Relatively Unaltered in Embryonic Hearts..... 46
3.3.5. 5hmC Loss Reduces YY1 Binding to Its Genomic Targets by Modulating Chromatin Accessibility in Embryonic Cardiac Tissues..... 48
3.3.6. 5hmC loss disrupts higher-order chromatin structures in embryonic hearts..... 52
3.4. Discussion..... 55
3.5. Materials and Methods..... 56
3.5.1. Human Samples..... 56
3.5.2. Animal Models..... 57
3.5.3. Antibodies..... 58
3.5.4. Histological Analysis..... 58
3.5.5. Quantification of the Compact Myocardium Thickness and the Trabecular Area..... 59

3.5.6. Immunohistochemistry (IHC) and Immunofluorescence (IF) 59
3.5.7. Nuclear Fractionation and Western Blot 60
3.5.8. Real-Time Quantitative PCR (qPCR) 60
3.5.9. RNA-Seq Library Construction and Data Analysis 62
3.5.10. Single cell RNA-seq (scRNA-seq) library preparation and data analysis 63
3.5.11. WGBS Library Construction and Data Analysis 64
3.5.12. ATAC-seq Library Construction and Data Analysis 65
3.5.13. Cleavage Under Targets and Release Using Nuclease (Cut&Run) 65
3.5.14. HiChIP Library Construction and Data Analysis 66
3.5.15. Prediction of AB Compartment 68

4. TET MEDIATED EPIGENETIC REGULATION IS IMPORTANT FOR SECOND HEART FIELD DEVELOPMENT 70

4.1. Summary 70
4.2. Introduction 70
4.3. Results 71
4.3.1. Complete Depletion of Tet Enzymes Caused the Second Heart Field Development Defects 71
4.3.2. Tet Triple Knock-Out Caused the Defects of Second Heart Field Progenitors During Heart Development 74
4.3.3. Tet Triple Knock-Out Affected the Epithelial to Mesenchymal Transition (EMT) of SHF Progenitors During Heart Development 77
4.3.4. EMT Related Transcriptional Factors Might Be the Target of Tet Enzymes 79
4.4. Discussion 80
4.5. Materials and Methods 81
4.5.1. Animal Models 81
4.5.2. Antibodies 82
4.5.3. Paraffin Section of Embryos 83
4.5.4. Hematoxylin-Eosin(H&E) Staining and Quantification 83
4.5.5. Immunofluorescence (IF) Staining 83
4.5.6. Real-Time Quantitative PCR (qPCR) 84
4.5.7. Single Cell RNA-Seq Libraries Preparation and Data Analysis 85
4.5.8. RNA-seq Libraries Preparation and Data Analysis 85

5. CONCLUSIONS 86

6. REFERENCES 88

LIST OF FIGURES

Page

Figure 1.1.DNA Methylation and Demethylation. 2

Figure 1.2.Tet Enzymes and Active DNA Demethylation. 3

Figure 1.3.Major Stages of Mouse Embryonic Heart Development. 9

Figure 2.1.Dynamic Tet and 5hmC Changes During Mesoderm-Cardiac Progenitor Differentiation of Nkx2-5-EmGFP mESCs. 13

Figure 2.2.Tet Depletion Compromises the Efficiency of mESC-to-Cardiac Progenitor Differentiation. 15

Figure 2.3.Single cell RNA-seq (scRNA-seq) Analysis Reveals Differentiation Defects During the mESC-to-Cardiac Progenitor Differentiation. 18

Figure 2.4.Genome-wide 5hmC Profiling During mESC-to-Cardiac Progenitor Differentiation. 21

Figure 2.5. Tet1-Mediated DNA Methylation Oxidation Directly Contributes to Cardiac Progenitor Differentiation. 23

Figure 3.1.Dynamic Changes of DNA Methylation and Hydroxymethylation During Embryonic Heart Development in Mmice and Human. 39

Figure 3.2.Cardiac-Specific Deletion of Tet2 and Tet3 Resulted in Developmental Defects in the Ventricular Chamber. 42

Figure 3.3.Transcriptomic Analyses on Control and Tet2/3-DKO Embryonic Heart Tissues. 45

Figure 3.4.Tet2/3 Deletion in Embryonic Heart Resulted Impaired 5hmC but not 5mC. 48

Figure 3.5.Tet2/3 Deletion in Murine Embryonic Hearts Reduced Chromatin Accessibility and Compromised the Binding of YY1 to Chromatin. 51

Figure 3.6.Tet-Mediated DNA Hydroxymethylation Is Associated with Chromatin Higher-Order Structures. 54

Figure 4.1.Complete Depletion of Tet Enzymes Caused the Second Heart Field Development Defects. 73

Figure 4.2.Tet Triple Knock-Out Blocks Second Heart Field Progenitors Differentiating to Myocytes. 76

Figure 4.3.Tet-TKO Affected the Epithelial-Mesenchymal Transition of SHF Progenitors. 78

Figure 4.4.EMT Related Transcriptional Factors Might Be the Target of Tet Enzymes. 80

LIST OF TABLES

Page
Table 2.1.Primers for Real-Time Quantitative PCR (Part1)..... 29

Table 3.1.Genotyping Primers of Tet, Tet3 and Nkx2.5Cre Mice 57

Table 3.2.Primers for Real-Time Quantitative PCR (part2)..... 60

Table 4.1.Genotyping primers of Tet1, EYFP and Isl1MerCreMer Mice..... 82

Table 4.2.Primers for Real-Time Quantitative PCR (Part3)..... 84

1. INTRODUCTION

1.1. DNA Methylation and Demethylation

In mammalian cells, 5-methylcytosine (5mC) is one of the major DNA epigenetic marks. It accounts for nearly 8% cytosines and often observed at CG dinucleotide position (termed CpG). DNA methylation counts for about 60% CpG sites in the genome[1,2]. It involves many important biological processes, such as genomic imprinting, X-chromosome inactivation, transcription silencing, and plays pivotal roles in normal development and diseases[3-5].

DNA methylation and demethylation balance is controlled by two family members, DNA methyltransferases family and TET dioxygenases family. The cytosine methylation is catalyzed by the DNA methyltransferases (DNMTs) (Figure 1.1), which catalyze methyl group at the 5-position of the pyrimidine ring of the cytosine and form the 5mC[6]. Among all the DNMTs family members, DNMT3a and DNMT3b are de novo DNMTs that enable DNA methylation at unmethylated cytosine. DNMT1 prefers to recognize the hemi-methylated CpG sites and maintained DNA methylation fidelity in the newly synthesized strand after DNA replication[7-10].

On the other hand, DNA demethylation is the process to replace 5mC with unmethylated cytosine using two different mechanisms: passive and active DNA demethylation. The passive DNA demethylation is due to loss of DNA methylation modifications during DNA replication, while the active DNA demethylation is mediated by the TET dioxygenases family[11-13].

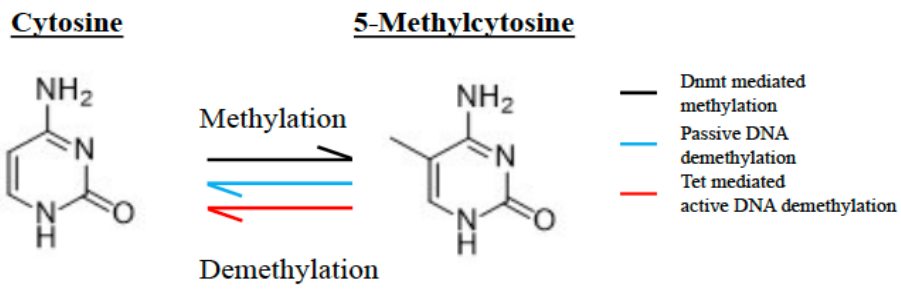

Figure 1.1.DNA Methylation and Demethylation.
Black line represents the Dnmt mediated DNA methylation; green line indicates the passive DNA demethylation; red line indicates the Tet mediated active DNA demethylation.

1.2. Tet Enzymes and Active DNA Demethylation

Tet-eleven translocation (Tet) proteins mediated active DNA demethylation was discovered in 2009 [14,15]. TET proteins belong to dioxygenase family which mediate the oxidation of 5mC to successively generate 5-hydroxymethylation (5hmC, the major product), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), ultimately causing DNA demethylation through base excision repair mechanism[16](Figure 1.2A). TET protein contains three members, TET1, TET2 and TET3, which are all catalytically active [17,18] and share similar domain structures (Figure1. 2B). The catalytic function was achieved by a double-stranded β-helix (DSBH) domain and a cysteine-rich domain. The catalytic activity of TET protein is dependent on iron (Fe (II)) and α-ketoglutarate (α-KG), where their binding sites are located within DSBH domain. The cysteine-rich domain is important to stabilize the interaction between Tet protein and DNA [19,20]. Tet1 and Tet3 contain CXXC domain at the N-terminus enabling the recognition of unmethylated CpG

sites, while the putative CXXC domain is separated with Tet2 during evolution and known as Idax [21].

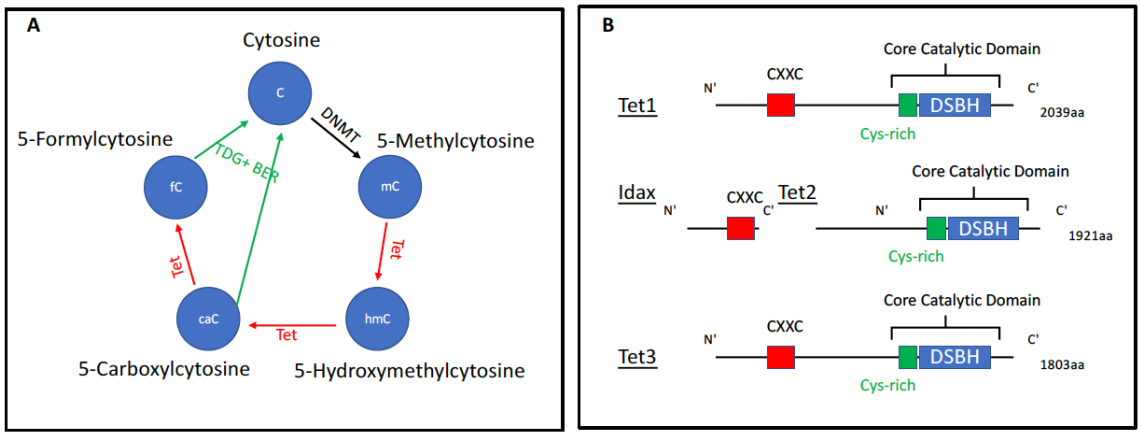

Figure 1.2.Tet Enzymes and Active DNA Demethylation.
A) Tet mediated active DNA demethylation; DNMT, DNA methyltransferases; TDG, thymine-DNA glycosylase; BER, base excision repair. B) Domain structure of mouse Tet enzymes; DSBH, double-stranded β-helix.

Tet mediated active DNA demethylation is tightly regulated. Its catalytic activity is heavily dependent on various co-factors and metabolites, including oxygen, α-KG, and reduced form of Iron [22-24,25]. It is reported that Vitamin C can enhance Tet catalytic activity by maintaining Iron in a reduced form [26,27]. Three metabolites, 2-hydroxyglutarate(2-HG), succinate and fumarate, have similar structures with α-KG and can competitively inhibit Tet enzymatic activity[28,29]. Gain of function mutations in metabolic enzymes isocitrate dehydrogenases 1 and 2 (IDH1/2) lead to aberrant accumulation of 2-HG, which decreased Tet enzymatic activity and cased the acute myeloid leukemia in human[24,28].

The transcription and protein level of TET proteins are also tightly regulated by various mechanisms. For example, microRNAs, such as miR-15b, miR-22, miR-26, miR-29, miR-

125, miR-494 and miR-302/367, are reported to reduce the transcription level of Tet [30-33]. Ubiquitylation, acetylation, phosphorylation, GlcNAcylation and PARylation of Tet proteins have been reported and affect the subcellular localization, chromatin binding and enzymatic activity of TET proteins[34]. Recent studies showed that Caspase-dependent degradation, calpain-mediated proteolysis or ubiquitin–proteasome pathway could regulate Tet2 protein level in different systems[35].

Tet mediated DNA methylation oxidation plays critical roles in various biological systems. For example, Tet2 deficiency resulted in increased hematopoietic stem cells and myeloid bias during hematopoiesis[36]. Tet3 deletion resulted in depletion of the neural stem cell pool *in vivo*[37]. More details will be elaborated in the following sections. At molecular level, Tet protein mediated epigenetic regulation is important for regulating chromatin accessibility and enhancer activity [38-41]. In addition, it is involved in regulating RNA splicing by gene-body demethylation [42-44]. Overall, TET mediated epigenetic regulation is important for transcriptional regulation at different levels.

1.3. Tet Enzymes Mediated DNA Demethylation in Embryonic Stem Cells

Embryonic stem cells (ESCs) are derived from the inner cell mass of a blastocyst, a very early embryonic development structure and maintained in an undifferentiated, pluripotent state. Stem cells can be differentiated toward different lineages with various growth factors and the dramatic epigenetic landscapes changes have been well documented during ESC differentiation[45]. Although deletion of TET proteins has minor effects on ESC pluripotency, TET protein mediate epigenetic regulation is involved in maintaining the expression of several key pluripotency factors. For example, Tet1 maintains low DNA methylation at Nanog promoter [46-48] in mouse ESC (mESC) and TET2 mediated DNA

demethylation enabling the transcription of NANOG in human ESCs (hESC) [49]. In addition, while Telomere is required for the self-renewal and pluripotency of ESCs. Tet mediated epigenetic regulation is important for transcription of telomere recombination genes and maintaining the telomeres length, which is important for ESC self-renewal and pluripotency [50,51].

Tet mediated DNA methylation oxidation is essential for ESC lineage specification. Tet1 controls the expression of trophectoderm and neuroectoderm differentiation related genes during mESCs differentiation. Tet1 deficient mESCs skews to differentiate into endoderm-mesoderm lineages[47]. Tet1 and Tet2 double deficient mESCs tend to differentiate to the endoderm and increase the extra-embryonic lineage differentiation[52-54]. Tet3 deficient mESCs have higher methylation level in secreted frizzled-related protein 4 (Sfrp4) and deceased expression of Sfrp4, which partially impair the differentiation of the neuroectoderm[55]. Tet1/2/3 triple deficient mESCs formed the poorly differentiated embryonic bodies and had server mesoderm differentiation defect[46]. Similarly, TET1/2/3 triple deficient hESCs also exhibited the poorly differentiated embryonic body and failed to form the teratomas. At molecular level, TET1/2/3 deficiency resulted in hypermethylation at bivalent promoters, such as PAX6 promoter, and failed to turn on this critical transcription factors during differentiation[43]. Taken together, Tet enzymes mediated epigenetic regulation is essential for lineage specification during ESCs differentiation.

1.4. Tet Enzymes Mediated DNA Demethylation in Mammalian Embryonic Development

Embryonic development initiates from the fertilization, which a haploid male gamete fuses with a haploid female ovum to generate the diploid zygote. The zygote then undergoes 2-cell, 4-cell, 8-cell and 16-cell stages to form the blastocyst[56]. During these stages, cells only perform mitotic divisions without cell size growth and differentiation. Before blastocyst stage, cells have the pluripotency to differentiate to any types of cells in the body[57]. The blastocyst then implanted to the wall of the uterus followed by the gastrulation to form the three germ layers including, endoderm (inner layer), mesoderm (middle layer) and ectoderm (outer layer). Gastrulation is a very crucial stage during early embryonic development[58]. Any development caveats at this stage will cause severe developmental defects. Embryo will perform the early organogenesis and cells from three germ layers will contribute to the development of different organs. The endoderm mainly contributes to the digesting system, the respiratory system, and other organs, such as the liver, the pancreas, the bladder and the thyroid. The mesoderm forms most parts of the cardiovascular system, the hematological system, and the genitourinary system. The ectoderm consists of the surface ectoderm, neural crest, and the neural tube[59-63].

The process of mammalian embryonic development is tightly controlled, and DNA methylation and demethylation play critical roles. Tet mediated active DNA demethylation is crucial for embryonic development[48,64]. Various Tet deficient mouse models were established to study the function of Tet proteins during embryonic development. Tet1 deficient mouse has a slightly smaller body size and minor defect in the reproductive system[65]. Tet2 deficient mice model exhibited increased hematopoietic

stem cells and myeloid bias [36,66,67]. Tet3 deficiency resulted in embryonic lethality and affected brain development[68]. Tet1 and Tet2 double knockout (DKO) mice showed defects in reproductive and hematopoietic system[69]. Tet1 and Tet3 DKO mice are detrimental to embryogenesis with developmental defects at 8- cell stage[70]. Tet2 and Tet3 DKO mice are also detrimental with the major defects found in the hematopoietic system[71]. Tet1/2/3 triple knockout (Tet-TKO) mice developed the lethal defects at the gastrulation stage with strong defects in mesoderm germ layer[72]. Altogether, findings revealed from the various Tet deficient mouse models suggest that Tet enzymes are required for normal embryonic development.

1.5. Tet Enzymes Mediated DNA Demethylation in Heart Development

To further evaluate tissue specific function of Tet protein, it is important to develop tissue specific Tet deficient mouse models. Heart is the first organ developed after gastrulation and is essential for embryonic development. Mouse embryonic heart development is a complicated process (Figure 1.3). It initiates from the first and second heart fields in the cardiac mesoderm at around E7.5. At this stage, cardiac progenitors from both heart fields will proliferate and differentiate to form the linear heart tube at around E8.5. This linear heart tube will start to loop at E9.5 and make the atria at the top at the ventricle at E10.5. After E10.5, the heart will continue to develop with the formations of atrial and ventricular septa, the compaction of myocardium, and the formation of vascular system to establish the well-known four-chambered organ[73-75].

Heart development is precisely controlled. Abnormal heart development leads to the congenital heart diseases, which is the most common birth defects in human. DNA methylation plays important roles in regulating heart development. Methylation levels in

the regulatory elements, such as promoter and enhancer, and gene bodies of the cardiac specific transcriptome are dynamically changed during heart development[76,77]. For example, DNA methylation in fetal cardiomyocyte genes, such as NPPA, TNNI1 and MYL4, is gradually increased during heart development; and DNA methylation in mature cardiomyocyte genes, such as MYH6, RYR2 and TNNI3, is gradually decreased during heart development[76]. Abnormal DNA methylation during heart development is associated with congenital heart diseases. Studies showed that DNA methylation status in hearts from congenital heart diseases patients is dramatically different with DNA methylation status in normal health hearts[78-80].

There are several studies reported the essential roles of Tet protein in cardiac lineage specification and development. Li et al. showed that both Tet3-KO and Tet-TKO mESCs had impaired neural conversion and Tet deficiency skewed the differentiation toward cardiac mesoderm[55]. Greco et al. showed that 5hmC was highly enriched in the bodies and distal regulatory elements of cardiac development related genes in E14.5 and neonatal hearts[44]. Lan et al. reported that Tet2 and Tet3 mediated DNA demethylation regulate the expression of genes important for endocardium and myocardium and control the epicardium cell migrating to the heart[81]. Taken together, we hypothesize that Tet enzymes mediated epigenetic regulation plays crucial roles in heart development by orchestrating the cardiac specific transcriptomic networks. However there lacks systematic analysis to reveal the function of Tet protein in cardiac development. Therefore, we would like to use mESCs cardiac differentiation model and various cardiac-specific Tet deficient mice

models to investigate the function of Tet protein in cardiac lineage specification and heart development.

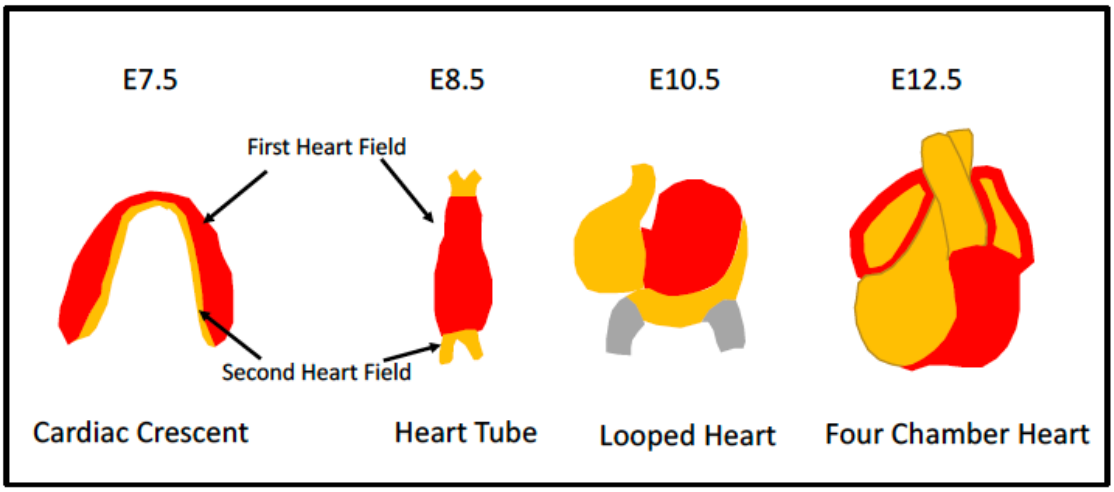

Figure 1.3.Major Stages of Mouse Embryonic Heart Development.
Mouse heart development starts from the cardiac crescent in the mesoderm at embryonic day 7.5. Cells from first and second heart fields form the heart tube. Heart tube starts to loop at embryonic day 8.5 and makes the atrium above the ventricle.

2. TET ABLATION IMPAIRS MOUSE EMBRYONIC STEM CELL CARDIAC DIFFERENTIATION

2.1. Summary

TET dioxygenases mediated DNA methylation oxidation plays an important role in regulating the embryonic stem cells (ESC) differentiation. In this part, we applied a CRISPR/Cas9 based genome editing method to generate single, double and triple Tet-deficient mouse ESCs (mESCs) and differentiated these cells toward cardiac lineage. By using emerald GFP (emGFP) expression under the control of Nkx2.5 promoter as marker for cardiac progenitor cells, we discovered that Tet1 and Tet2 depletion, as well as Tet triple depletion, significantly impaired mESC-to-cardiac progenitor differentiation. Single-cell RNA-seq analysis further revealed that Tet deletion resulted in the accumulation of mesoderm progenitors to hamper cardiac differentiation. Re-expression of the Tet1 catalytic domain (Tet1CD) rescued the differentiation defect in Tet-TKO mESCs. dCas9-Tet1CD mediated loci-specific epigenome editing at the Hand1 loci validated the direct involvement of Tet-mediated epigenetic modifications in transcriptional regulation during cardiac differentiation. This study establishes that Tet-mediated epigenomic remodeling is essential for maintaining proper transcriptional outputs to safeguard mESC-to-cardiac progenitor differentiation.

2.2. Introduction

Tet enzymes mediated DNA demethylation is one of the epigenetic factors that ensure the proper transcriptional outputs and lineage commitment during development[46,69,72]. Embryonic stem cells (ESCs) differentiation model is useful for investigating the

pathways in early development and cellular lineage specification[43,55,82]. Previous studies showed that Tet enzymes mediated DNA demethylation plays crucial roles in ESCs self-renewal and maintaining pluripotency[42,83,84]. However, the function of Tet proteins in regulating mESC differentiation toward cardiac lineage is still unclear.

In this study, we generated single, double and triple Tet enzymes knock-out E14 mESCs expressing emerald GFP (emGFP) under the Nkx2.5 promoter, by using CRISPR/Cas9 based genome editing[85,86]. Hanging-drop differentiation method was performed to induce WT and Tet deficient mESCs towards cardiac lineage[87] . We observed that deletion of Tet1 and Tet2 significantly impaired the differentiation of mESC toward cardiac progenitors. Single-cell RNA-seq (scRNA-seq) analysis in Tet-triple knockout (Tet-TKO) cells revealed accumulation of cardiac mesoderm progenitor cells during differentiation, which might impede the proper cardiac progenitor lineage specification. DNA methylation and hydroxymethylation analysis further confirmed that Tet-mediated DNA methylation oxidation is critical for ensuring the proper transcriptional regulation during mESC-to-cardiac progenitor differentiation. Our studies demonstrated that Tet protein mediated DNA demethylation is one of the critical epigenetic regulators for proper cardiac lineage specification.

2.3. Results

2.3.1. Dynamic DNA hydroxymethylation changes during mESC-to-cardiac progenitor differentiation.

To investigate how Tet-mediated DNA methylation oxidation regulates cardiac progenitor differentiation, we used a hanging drop differentiation procedure for in vitro cardiac progenitor differentiation (Figure 2.1A). To conveniently monitor cardiac progenitor differentiation efficiency, we used mESCs expressing emGFP under the Nkx2.5 promoter

(Nkx2-5-EmGFP mESCs) as our differentiation model system. We detected the change of global 5hmC levels during cardiac progenitor differentiation using dot-blot assay. We observed an initial decrease of 5hmC within the first 4 days of differentiation, followed by a gradual increase of 5hmC from Day 6 to Day 10 during this differentiation process (Figure 2.1B). In parallel, we performed real-time quantitative PCR (qPCR) to monitor the transcription of three Tet genes during this differentiation process. Consistent with the dot-blot analysis, we observed pronounced decrease of Tet1 and Tet2 expression within the first two days of differentiation (Figure 2.2B). During Day 4 and Day 10, we observed a gradual upregulation of both Tet2 and Tet3 during mESC-to-cardiac progenitor lineage specification (Figure 2.1C), suggesting that Tet2 and Tet3 might be responsible for increased 5hmC during in vitro differentiation.

To further pinpoint which Tet proteins are important for cardiac progenitor lineage specification, we generated single, double and triple Tet-deficient Nkx2-5-EmGFP mESCs using a CRISPR/Cas9 based genome editing method. We observed that single deletion of Tet had minor effects on cardiac progenitor differentiation; while both Tet1/2-double and Tet1/2/3-triple deficient mESCs exhibited significant reduction in emGFP levels during mESC-to-cardiac progenitor differentiation, suggesting that Tet1 and Tet2 are the major players in regulating in vitro mESC differentiation toward cardiac progenitors (Figure 2.1D).

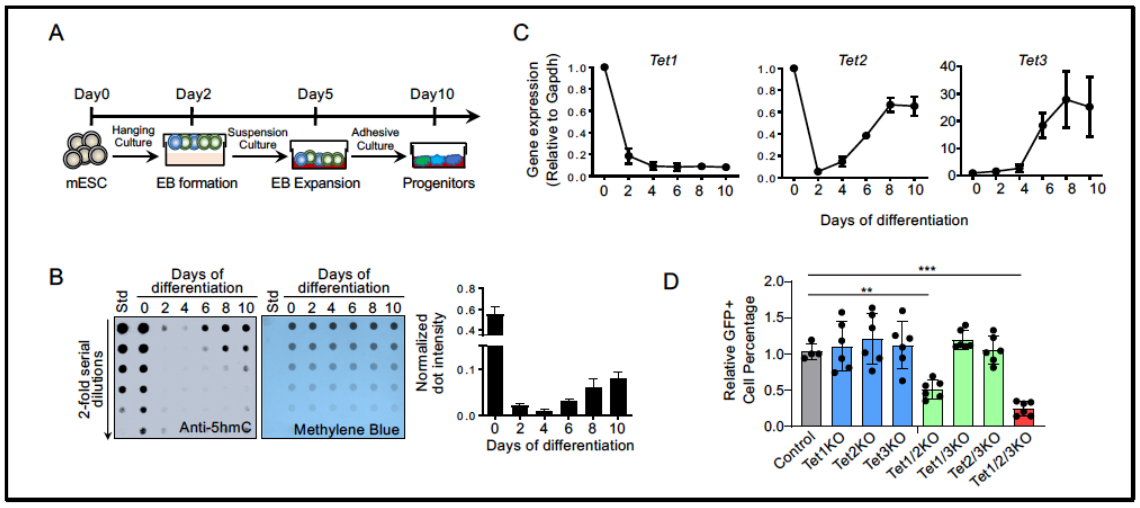

Figure 2.1.Dynamic Tet and 5hmC Changes During Mesoderm-Cardiac Progenitor Differentiation of Nkx2-5-EmGFP mESCs.

A) Schematic illustrating in vitro hanging drop differentiation procedures; B) Dot-blot analysis of 5hmC levels during mESC-to-cardiac progenitor differentiation; C) Real-time qPCR analysis of Tet1, Tet2, and Tet3 expression at different days during Nkx2-5-EmGFP mESC cardiac differentiation. Data were shown as mean ± S.D; n=3; D) mESC-to-cardiac progenitors differentiation efficiency measured by GFP+ cells expressing Nkx2.5 at day 5 of directed differentiation. Data were shown as mean ± S.D; n=3 individual clones with two biological replicates. ** $p < 0.01$, ***$p < 0.001$(compared to Control, one-way ANOVA with Tukey post-hoc test).

2.3.2. Tet Triple Knockout (TKO) mESC Exhibits Defects During Its Differentiation Toward Cardiac Progenitors.

To avoid potential compensatory effects of Tet proteins, we next chose two clones of Tet triple knockout (TKO) mESC for in vitro cardiac progenitor differentiation. We measured the emGFP levels during this differentiation process. During day 0 to day 10 differentiation, we observed GFP signal reached the highest level at day5 and a notable

reduction in the emGFP-positive population in both TKO clones when compared with the control group (Figure 2.2A). Immunoblot analysis further revealed a reduction in the expression of cardiomyocytes markers, α-actinin and troponin T (cTnT), in Tet-TKO clones compared with the control group after 10 days of cardiac differentiation (Figure 2.2B). Therefore, we focused on analyzing these progenitors at day 5 using the hanging drop differentiation method. We observed ~75% decreased the embryonic body (EB) size in Tet-TKO mESCs when compared to the control group at day 5 of differentiation (Figure 2.2C, E). Furthermore, we observed that the control group displayed a relatively flat shape with high beating capability (Figure 2.2D, F). In contrast, Tet-TKO cells exhibited a round morphology with lower beating capability after 10 days differentiation (Figure 2.2D). In parallel, we examined the expression of pluripotent and lineage markers in both the control and Tet-TKO groups at day 0 and day 5 of differentiation (Figure 2.2G). We observed upregulation of trophectoderm (Eomes), endoderm (Sox17), mesoderm (Brachyury) markers, but significant downregulation of cardiac mesoderm markers, such as Nkx2.5 and Tbx20, in the Tet-TKO group at day 5 of differentiation. Collectively, these data suggest that Tet deletion impairs cardiac progenitor differentiation of mESCs.

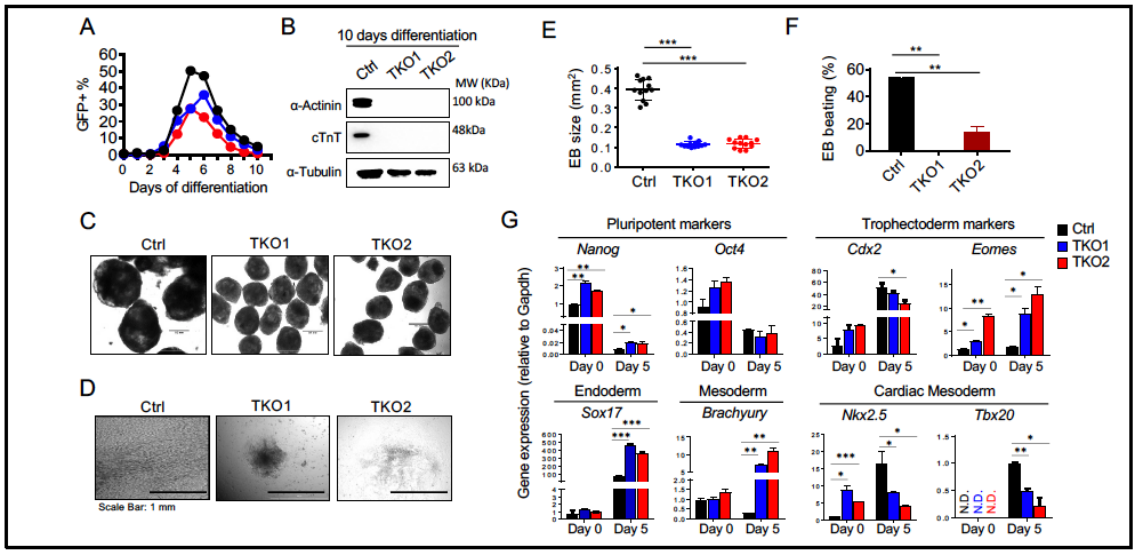

Figure 2.2.Tet Depletion Compromises the Efficiency of mESC-to-Cardiac Progenitor Differentiation.

A) Quantification of the percentage of GFP-positive cells in the control and TKO groups (two clones) at different days during mESC-to-cardiac progenitor differentiation. Data were shown as mean ± S.D; n= 3; B) Immunoblot analysis of actinin and cTnT expression in control and TKO mESCs differentiated for 10 days. α-tubulin was used as loading control; C) Representative images (4X) of embryonic bodies (EBs), Scale bar: 300 µm; D) Representative images of differentiated cardiac progenitor clones derived from the control and TKO mESCs (two clones) at 10 days of differentiation, Scale bar: 1 mm.; E) Quantification of the EB size in the control and TKO groups (clone 1 and clone 2). Data were shown as mean ± S.D; n= 60 EBs. *** $p < 0.001$ (compared to the control), by two-tailed Student's t-test; F) Quantification of the percentage of beating clones in control and TKO (two clones) groups at 10 days of differentiation. Data were shown as mean ± S.D; n= 3. ** $p < 0.01$ (compared to the control; two-tailed Student's t-test); G) Real-time qPCR analysis of gene expression in control and TKO (two clones) Nkx2-5-EmGFP mESCs at

day 0 and day 5 of differentiation. Data were shown as mean ± S.D; n= 2. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ (compared to the control; two-tailed Student's t-test).

2.3.3. Single-cell RNA-seq (scRNA-seq) Analysis Reveals Abnormal Cardiac Progenitor Lineage Specification in Tet-TKO mESCs.

To further portrait the molecular features of Tet-TKO mESC during differentiation, we performed single-cell RNA-seq (scRNA-seq) analysis on differentiation day 5 control and Tet-TKO mESCs. We profiled a total of 7,539 cells with 4,381 cells in the control group and 3,158 in the TKO group. After unbiased clustering on aggregated cells using Seurat[88], we identified 6 distinct clusters in all the analyzed cells (Figure 2.3A, B). Based on the gene expression signatures identified from each cluster, most clusters were mesoderm cells or mesoderm derived progenitor cells for different lineages, including hematologic, mesenchymal, cardiac progenitors (Figure 2.3A-B). For example, cluster 2 exhibited high expression of transcription factors (TFs), Isl1, Tbx3 and Gata3, which are critical for early embryonic development and indicative of a mesoderm-progenitor-like feature of this cluster.

We further compared each cluster between the control and Tet-TKO groups and observed decreased cell numbers in the Tet-TKO group in clusters with mesoderm-like and mesoderm-derived progenitors-like features (including clusters 0, 1, 3, 4) compared with the control group (Figure 2.3C). Cardiac cells in Tet-TKO were slightly increased (<2% of total cells) and this might be due to low cell number and large variation since majority cells at this stage are mesoderm cells. Indeed, cell numbers in cluster 2 with a mesoderm progenitor-like feature are significant increase (Figure2.3C) and this finding is confirmed by our RNA velocity analysis[89] (Figure2.3D). These results suggest that Tet deletion

blocks the mesoderm differentiation toward cardiac progenitors which is consistent with real-time qPCR and western blotting analysis (Figure 2.2B, G).

Next, we performed pseudo time trajectory analysis[90] to identify key genes that might contribute to the difference between control and Tet-TKO group during differentiation. We identified three major branch points that were regulated by Tet protein during this differentiation process (Figure2.3E). Within the three branch points, the control and Tet-TKO groups exhibited distinct distribution at branch A, but showed overlapping distribution at branches B and C, suggesting that branch A might be critical for Tet-regulated cardiac differentiation. Key mesoderm development regulators, including Hand1, Bmp4 and Snai2, are in branch A and their expression were significantly downregulated in Tet-TKO groups (Figure 2.3F, G). The disruption of the expression of these genes might contribute to the impaired cardiac progenitor lineage specification in Tet-TKO mESCs[91-93]. Overall, the scRNA-analysis establishes that deletion of Tet in mESCs might compromise cardiac progenitor specification through blocking the differentiation of mesoderm progenitors toward the downstream lineages.

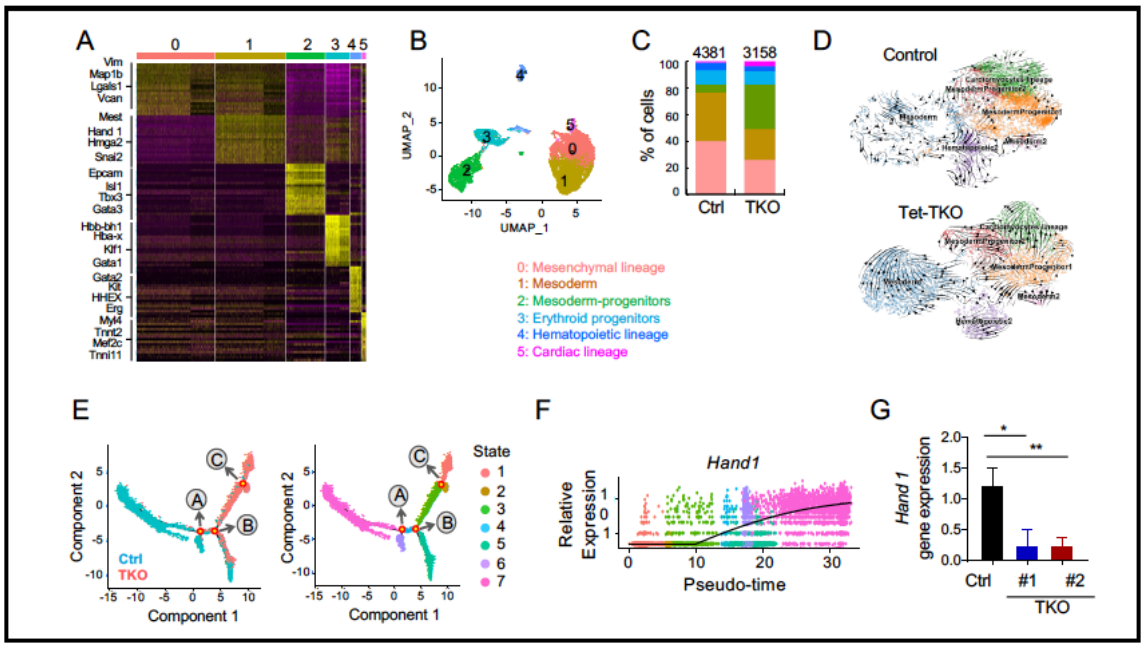

Figure 2.3.Single cell RNA-seq (scRNA-seq) Analysis Reveals Differentiation Defects During the mESC-to-Cardiac Progenitor Differentiation.

A) Heatmap representation of expression signatures for each of clusters in all analyzed cells; B) UMAP dimensionality reduction (n = 7,539 cells) in all analyzed cells that were collected from day 5 of differentiation; C) Quantification of the percentage of cell types in control and TKO mESCs when differentiated toward cardiac progenitors for 5 days. The total analyzed cell numbers were shown above the bars; D) RNA-velocity analysis reveals the differentiation process projected in UMAP-based embedding in the control and Tet-TKO mESCs; E) Pseudotime trajectory analysis of scRNA-seq data obtained from the control (cyan) and TKO (red) mESCs after directed differentiation toward cardiac progenitors for 5 days. Branches labeled with A, B, C represent the points that are critical for separating different differentiation status as indicated on the right using variegated colors; F) Dynamic changes of Hand1 expression during mESC-to-cardiac progenitor differentiation based on pseudotime trajectory analysis. The solid line indicates the mean

expression level of Hand1 at different differentiation status; G) Real-time qPCR analysis of Hand1 expression in the control and TKO groups (two clones) at day 5 of differentiation. Data were shown as mean ± S.D; n= 3. ** $p < 0.01$, * $p < 0.05$ (compared to the control; two-tailed Student's t-test).

2.3.4. Tet Deficiency Impairs DNA Methylation at the Hand1 Promoter.

To examine how Tet protein mediated DNA methylation oxidation contributes to transcriptional regulation during cardiac progenitor differentiation, we performed genome-wide 5-hydroxymethylation analysis using the CMS-IP-seq method[94] in genomic DNA collected from mESCs after differentiation toward cardiac progenitors for 0, 2 and 5 days. Consistent with the dot-blot analysis results, we observed dynamic 5hmC changes during this differentiation process (Figure 2.4A). We then performed pair-wise comparison and identified differential 5hmC enriched regions (DHMRs) at different days. Genomic Regions Enrichment of Annotations Tool (GREAT) analysis[95] of these DHMRs showed that stage-specific 5hmC peaks were enriched at potential regulatory regions for genes that are important for mesoderm and cardiac development (Figure 2.4B). In parallel, we observed that the increased DHMRs at day 2 were significantly associated with transcription factors (TFs) that are essential for cardiac development, such as Tbx5 and Gata4[96,97]; while the decreased DHMRs at day 2 were significantly enriched with motifs for TFs, such as Esrrb and the ETS family, which are important for stemness and neuronal differentiation[98,99]. Similarly, genomic regions that displayed augmented DNA hydroxymethylation at day 5 of differentiation were enriched in TFs binding motifs for Gata4, Tbx5 and Isl1, which are important for cardiac lineage specification. These data

strongly suggest that Tet and the accompanying DNA hydroxymethylation modulate mESC differentiation toward cardiac progenitors.

Since scRNA-seq analysis indicated that Hand1 is among the most affected genes in Tet deficient mESCs, we set out to examine the 5hmC level at the Hand 1 locus in control mESC at different days during differentiation. We observed significantly increased 5hmC level within the Hand1 locus at day 2 and day 5 of differentiation compared with undifferentiated (day 0) mESCs (Figure 2.4C), suggesting that Tet protein mediated DNA methylation oxidation might directly regulate Hand 1 transcription. Next, we performed loci-specific DNA methylation analysis at Hand1 promoter and enhancers in the control and Tet-TKO mESCs that were differentiated for 5 days. Among all the tested loci, we pronounced increase of DNA methylation at the Hand1 promoter and slightly increased DNA methylation at a potential enhancer of Hand1 (Figure 2.4D). These data suggest that the impaired Hand1 gene expression in Tet-TKO cells might be due to the increased DNA methylation at its promoters and enhancers.

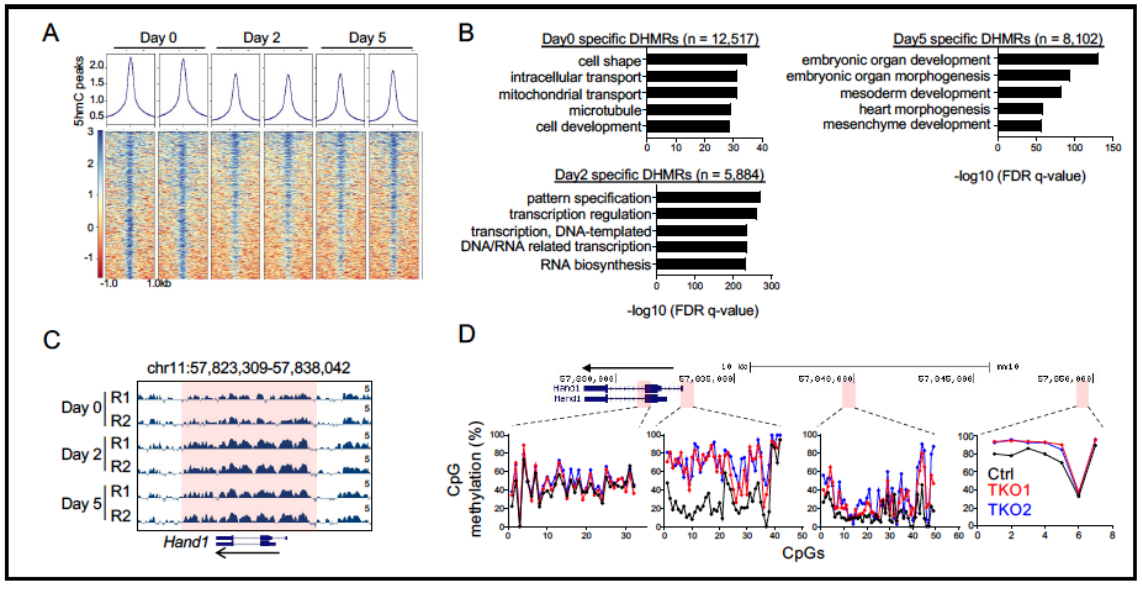

Figure 2.4.Genome-wide 5hmC Profiling During mESC-to-Cardiac Progenitor Differentiation.

A) The enrichment of 5hmC in the control mESCs at the indicated time points during mESC-to-cardiac progenitor differentiation; B) GREAT analysis of differentially hydroxymethylated regions (DHMRs) at the indicated comparison groups; C) The UCSC genome browser view of 5hmC enrichment at the Hand1 locus in the control group at day 0, day 2 and day 5 of differentiation toward cardiac progenitors. R1 and R2 indicate two biological repeats. The highlighted region indicates the Hand1 locus; D) Loci-specific DNA methylation analysis in Hand1 enhancers, promoter, and gene-body in the control (black) and TKO groups (two clones; red and blue) at Day 5 of differentiation. The analyzed regions were highlighted in red. Each dot indicates sequenced CpG sites. Each CpG site was covered by at least 100 times.

2.3.5. Tet1 Is Required for Hand1 Expression During Cardiac Progenitor Differentiation.

To probe whether Tet mediated DNA methylation oxidation is required for cardiac progenitor lineage specification, we introduced Flag-Tet1 catalytic domain (CD) into Tet-TKO mESCs to restore the 5hmC in the genome. We observed a remarkable increase of 5hmC production in Tet-TKO mESCs in day 0 and day 5 during differentiation (Figure2.5A). We then differentiated Tet-TKO mESCs expressing Flag-Tet1CD toward cardiac progenitors. Flag-Tet1CD re-expression in Tet-TKO mESCs were found to increase the EB size at day 5 during the differentiation (Figure 2.5B). Furthermore, we observed increase of cardiac progenitors marked by emGFP, as well as the beating ability of cardiomyocytes (Figure 2.5C, D). In addition, the Hand1 expression level was significantly upregulated with decreased DNA methylation at its promoter (Figure 2.5E,

F). These findings suggest that Tet mediated active DNA demethylation is important for cardiac progenitor differentiation.

To clarify whether Tet mediated DNA demethylation could directly control Hand1 gene expression, we subsequently introduced catalytic dead Cas9 (dCas9) fused with Tet1CD (dCas9-Tet1CD)[100] , together with sgRNA targeted to the Hand1 promoter. No significant global upregulation of 5hmC levels was noted in control and TKO mESCs following dCas9-Tet1CD expression, suggesting the high specificity of the epigenome editing approach (data not shown). We observed significant upregulation of Hand1 gene expression only in TKO mESCs expressing both dCas9-Tet1CD and sgRNAs (Figure 2.GF), indicating that Tet mediated DNA demethylation might directly contribute to the transcriptional regulation of Hand1. These data confirm that Tet protein mediated DNA methylation oxidation is essential for regulating transcriptional network during cardiac progenitor differentiation mESCs.

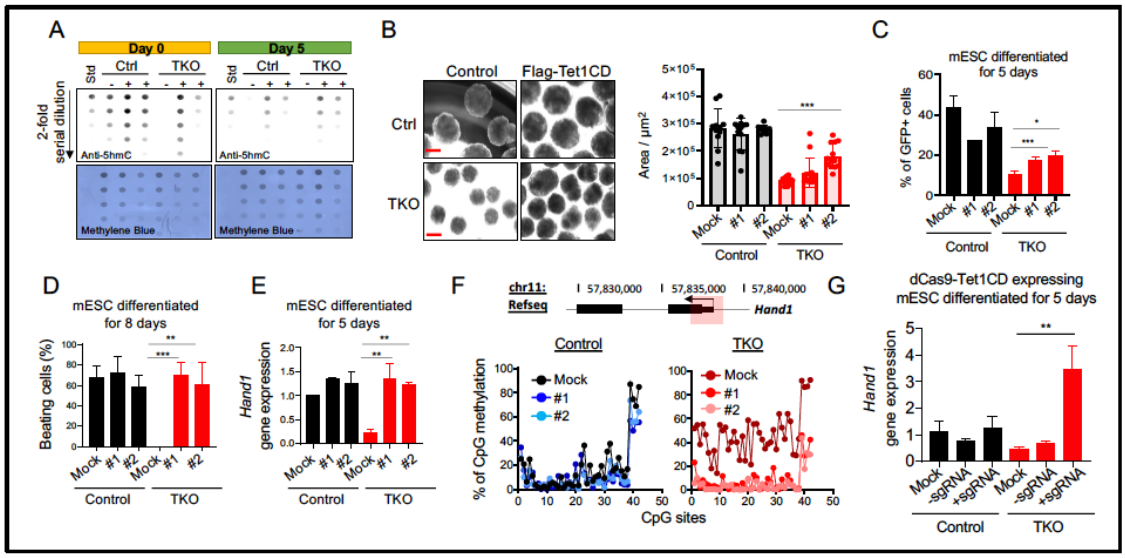

Figure 2.5. Tet1-Mediated DNA Methylation Oxidation Directly Contributes to Cardiac Progenitor Differentiation.

A) Dot-blot analysis of 5hmC measurement in control and TKO mESCs expressing Flag-Tet1CD at day 0 and day 5 of differentiation. The loading control was shown at the bottom of each blot using methylene blue to visualize the total amounts of input DNA; B) (Left) Representative images (4X) of embryonic bodies (EBs) under the indicated conditions. Scale bar: 300 μm. (Right) Quantification of the EB size in the control and TKO mESCs expressing Flag-Tet1CD (#1 and #2: two individual clones expressing Flag-Tet1CD). Data were shown as mean ± S.D; n= 12 EBs. *** $p < 0.001$(compared to mock; one-way ANOVA with Tukey post-hoc test); C) The percentage of GFP-positive (left) and Flk+Pdgfrα+ double positive (right) cells in the control and TKO mESCs expressing Flag-Tet1CD after 5-day differentation. Data were shown as mean ± S.D; n= 15 EBs. *** $p < 0.001$, * $p < 0.05$ (compared to mock; one-way ANOVA with Tukey post-hoc test); D) Quantificatin of the percentage of beating cells in the control and TKO mESCs expressing Flag-Tet1CD differentiated for 8 days. Data were shown as mean ± S.D; n= 60 EBs, ***$p < 0.001$, **$p<0.01$ (compared to mock; one-way ANOVA with Tukey post-hoc test); E) The expression of Hand1 in the control and TKO mESCs expressing Flag-Tet1CD differentiated for 5 days. Data were shown as mean ± S.D; n=3. ** $p < 0.01$ (compared to mock; one-way ANOVA with Tukey post-hoc test); F) Loci-specific DNA methylation analysis in the Hand1 promoter (highlighted in red) in the control (left) and TKO (right) mESCs expressing Flag-Tet1CD differentiated for 5 days. Each dot indicates sequenced CpG sites. Each CpG site was covered for at least 100 times; G) The expression of Hand1 in the control and TKO mESCs expressing dCas9-Tet1CD with or without sgRNA

targeted to the Hand1 promoter at Day 5 of differentiation. Data were shown as mean ± S.D.; n=3. ** $p < 0.01$ (compared to mock; one-way ANOVA with Tukey post-hoc test).

2.4. Discussions

In this study, we performed the in vitro cardiac progenitor differentiation using control and Tet-TKO Nkx2-5-EmGFP mESCs to investigate the role of Tet mediated DNA methylation oxidation during cardiac lineage specification[101]. We observed dynamic changes of 5hmC distribution during this differentiation process. Our study also reveals that deletion of Tet proteins significantly impairs cardiac progenitor differentiation. Based on our single-cell RNA-seq analysis, we observed that Hand 1 is one of the most affected genes in Tet-TKO mESC during this differentiation process. In parallel, we observed substantial increase of DNA methylation at the Hand1 promoters in differentiated Tet-TKO cells, suggesting that Tet protein mediated DNA methylation oxidation contributes to Hand1 expression. By using dCas9-based epigenome editing tool, we further validated the causal relationship between Tet protein mediated DNA methylation oxidation and the expression of genes (such as Hand1) that are essential for cardiac function. In addition, the differentiation defects imposed by Tet-TKO could be rescued by simple re-expression of the catalytic domain of Tet1, suggesting that Tet protein mediated epigenetic regulation is a genome-wide event and is required to coordinate the proper transcriptional outputs.

Although both Tet2 and Tet3 are upregulated during mESC-to- cardiac progenitor differentiation, our systematic CRISPR-aided genetic depletion studies reveal that Tet1 and Tet2 plays a more important role in regulating this differentiation process. Depleting Tet1 or Tet2 alone in mESC has minor effects on mESC differentiation, while deletion of Tet1 and Tet2 impaired cardiac differentiation, suggesting the compensation effects of

Tet1 and Tet2 during this differentiation process. Indeed, our rescue experiments using re-expression of the Tet1 catalytic domain have demonstrated the feasibility of using one Tet protein to restore the differentiation efficiency of Tet-TKO cells. Deletion of Tet3 could result in imbalanced differentiation between neuroectoderm and mesoderm[55]. Therefore, Tet3 might be important for the specification of other lineages other than cardiac progenitors. This is consistent with a previous study showing that Tet3-deficient mESC exhibits differentiation defects during neural progenitor differentiation[55].

Our study reveals that Tet protein mediated DNA demethylation directly contributes to Hand1 expression during cardiac differentiation. We also observed that other genes, such as Bmp4, exhibited significant transcriptional changes without DNA methylation alteration during cardiac differentiation (data not shown). It is likely that Bmp4 transcription is regulated by the non-catalytic functions of Tet proteins as reported previously[102], a direction warrants further investigation in follow-on studies.

In summary, our studies highlight a critical role of Tet protein mediated DNA methylation oxidation in mESC differentiation toward cardiac progenitors. Loci-specific epigenome editing using dCas9-Tet1CD to rescue single-gene expression could restore the DNA demethylation and transcription of targeted genes in Tet-TKO mESCs. Our findings suggest that a Tet protein mediated genome-wide epigenome remodeling is needed to coordinate the proper transcriptional outputs during mESC differentiation toward cardiac progenitors.

2.5. Materials and Methods

2.5.1. Mouse Embryonic Stem Cell (mESC) Culture and Differentiation

E14-Nkx2-5-EmGFP mESC was purchased from the Mutant Mouse Resource & Research Centers (MMRRC). Cells were cultured on mitomycin C treated mouse embryonic fibroblasts (MEF) as feeder cells. The culture medium includes Knock-out Dulbecco's Modified Eagle's Medium (Gibco), 15% fetal bovine serum (Omega), 0.5% penicillin-streptomycin (Gibco), 0.1 mM non-essential amino acids (Gibco), 0.1mM 2-mercaptoetanol (Sigma), and 103 U/mL of leukemia inhibitory factor (Millipore).

The hanging drop differentiation in mESCs was conducted according to a previous protocol. Briefly, mESCs were suspended in differentiation culture medium without leukemia inhibitory factor (LIF). Embryonic bodies (EBs) were formed by culturing cells within drops on the up-turned inner surface of the lid of the dishes for 2 days. EBs were then collected from the lids and cultured in non-adhesive dishes and subjected to further expansion for three additional days. The expanded EBs were cultured in the cell culture dishes to allow attachment for five additional days.

To measure the EB size after 5 days handing drop culture, the EBs were collected into a 15 ml conical tube and washed with PBS three times. EBs were then transferred to 96-well plates for further imaging analysis using Cytation5 (BioTek). Image J was then used to quantify the EB sizes. A polygon selection tool was applied to measure the EB size. To count the beating efficiency of EBs, individual EB were transferred to individual wells in 96-well plates. The beating phenotype was evaluated under a microscope.

2.5.2. Antibodies

Anti-flag (Sigma F1804, 1:2,000), anti-dCas9 (Cell Signaling Technology 7A9-3A3, 1: 1,000), anti-a-actinin(Abcam ab68167, 1:1,000), anit-cTnT(Theromo Scientific 13-11,

1:200), anti-Tublin(Santa Cruz sc-32293, 1:1,000), anti-MEST(Abcam ab95453, 1:1,000),anti-Tet1(gift from Dr. Margaret A Goodell lab in Baylor College of Medicine), anti-Tet2(Abcam ab124297), and anti-Gapdh (Sigma G8795, 1:3,000) were used for Western blotting. For dot-blot assays, an anti-5hmC (Active Motif 39769, 1: 40,000) was used. Biotin anti-mouse CD309 (Biolegend Clone 89B3A5, 1:100), PE anti-mouse CD 140a (Biolegend APA5, 1:100), Brilliant Violet421TM Streptavidin (Biolegend, 1: 400), and APC Streptavidin (Biolegend, 1:400) were used for FACS analysis.

2.5.3. Generating Tet Knockout mESCs

Tet knock-out mESCs were generated by the CRISPR/Cas9 technology as described previously with slight modifications. PX458 plasmids with validated sgRNAs targeting Tet1, Tet2, or Tet3 loci were transfected simultaneously into mESCs using the iMfectin DNA transfection reagent (GenDepot). GFP positive cells expressing Cas9 and sgRNA were sorted into 96-well plates by flow cytometry and individual colonies were genotyped after culturing for 7 days. Restriction Fragment Length Polymorphism (RFLP) analysis was performed to validate sgRNA targeting efficiency. Briefly, sgRNAs targeted regions were amplified by PCR and the products were digested with SacI (NEB), EcoRV (NEB) or XhoI (NEB), respectively. Digested PCR fragments were separated on a 2% SYBR Safe stained agarose gel (GenDepot). GFP, Cas9 and sgRNA expression was not detectable after positive clones were identified and used for cardiac differentiation experiments.

2.5.4. Establishing Flag-Tet1CD or dCas9-Tet1CD Expressing mESCs

Flag-Tet1CD was subcloned into a 213-pRRL lentivirus vector (a gift from Dr. Courtney Hodges at Baylor College of Medicine). The FUW-dCas9-Tet1CD plasmid was purchased from Addgene (#84475). sgRNAs targeting the Hand1 promoter (Table S2) were inserted

into the plko5.EFS.RFP647 plasmid (a gift from Dr. Margaret Goodell at Baylor College of Medicine). The lentiviruses encoding Flag-Tet1CD or dCas9-Tet1CD were produced via co-transfection with pSPAX2 (Addgene, #12260) and pMD2.G (Addgene, #12259) into HEK293T cells. Concentrated lentiviruses encoding Flag-Tet1CD or dCas9-Tet1CD were infected to control and Tet-TKO mESC. Polybrene (Sigma) was added to improve the infection efficiency. After 48 hours infection, single mESC were sorted into 96-well plates and individual clones were used for further characterization.

2.5.5. Genomic DNA, Total RNA and Protein Isolation

AllPrep DNA/RNA Mini Kit (Qiagen) was used to purify DNA and RNA. Purified DNA and RNA were measured by the Qubit fluorometer (Thermo Fisher Scientific). For protein extraction, cell pellets were washed with PBS twice and lysed in a RIPA buffer (150mM NaCl, 1% TritonX-100, 0.1% SDS, 0.5% Sodium deoxycholate, 50mM Tris-HCl pH8.0) on ice for 15 mins. Cell debris were removed by centrifugation at 4°C for 10 min (21,000 x g).

2.5.6. 5hmC Dot-Blot Assay

5hmC dot-blot assay was performed as previously described[103]. Briefly, Two-fold serial dilution was conducted to dilute the genomic DNA and the synthetic 5hmC standard oligo. DNA was denatured in 0.4 M NaOH, 10 mM EDTA at 95°C for 10 min, and then neutralized by ice-cold 2 M ammonium acetate (pH 7.0). The denatured DNA was hybridized on the nitrocellulose membrane by using a Bio-Dot apparatus (Bio-Rad). The membrane was vacuum baked at 80°C for 2 hours. After blocking with 5% non-fat milk for 1 hour at room temperature, the membrane was incubated with an anti-5hmC antibody (1:3000, Active Motif) overnight at 4°C. On the next day, the membrane was incubated

with secondary antibody (1:10,000; Sigma). West-Q Pico Dura ECL Solution (GenDepot) was used to develop the membrane and ImageLab (Bio-Rad) was used to visualize the 5hmC signal. After enhanced chemiluminescence treatment, the membrane was washed with 1X TBST briefly and then stained with 0.02% methylene blue in 0.3 M sodium acetate (pH 5.2) to confirm the total amounts of loaded DNA samples.

2.5.7. Real-Time Quantitative PCR (qPCR) Assay

ABScript II 1st Strand Kit (Abclonal) was used to generate cDNA. Real-time quantitative PCR was performed with a ViiA 7 instrument (Applied Biosystems) using a Universal SYBR Green Fast qPCR Kit (Abclonal). A three-step cycling program was used at 95˚C for 3 min, 40 cycles of 95˚C for 10 s, 60˚C for 20 s and 72˚C for 30 s. qPCR primers were synthesized by the Integrated DNA Technologies and the sequences were listed in Table 2.1.

Table 2.1.Primers for Real-Time Quantitative PCR (Part1)

Target	Direction	Sequences (5'->3')
Tet1	Forward	GAGCCTGTTCCTCGATGTGG
	Reverse	CAAACCCACCTGAGGCTGTT
Tet2	Forward	TGCAAAACCTGGCTACTGTC
	Reverse	AACATGCAGTGACTCCTGAG
Tet3	Forward	TCCGGATTGAGAAGGTCATC
	Reverse	CCAGGCCAGGATCAAGATAA
Nanog	Forward	AAGCAGAAGATGCGGACTGT
	Reverse	ATCTGCTGGAGGCTGAGGTA

Table 2.1. Primers for Real-Time Quantitative PCR (Part1) (Continued)

Target	Direction	Sequences (5'->3')
Oct4	Forward	TCTTTCCACCAGGCCCCCGGC
	Reverse	TGCGGGCGGACATGGGGAGA
Cdx2	Forward	AGGCTGAGCCATGAGGAGTA
	Reverse	CGAGGTCCATAATTCCACTCA
Eomes	Forward	TGCAAGAGAAAGCGCCTGTCTC
	Reverse	CAATCCAGCACCTTGAACGACC
Sox17	Forward	AAGAAACCCTAAACACAAACA
	Reverse	TTTGTGGGAAGTGGGATCAAG
Brachyury	Forward	GCATGCTGCCTGTGAGTCATA
	Reverse	ACCATTGCTCACAGACCAGAG
Tbx20	Forward	TATTCAGCATACTCCTAC
	Reverse	GTTAGTCTTGTCAATACG
Hand1	Forward	GGAGTTGCCTCAGCAGCCCG
	Reverse	CCGTAGCCGCTGCGTCCTTT
Gapdh	Forward	GTGTTCCTACCCCCAATGTGT
	Reverse	ATTGTCATACCAGGAAATGAGCTT

2.5.8. Western Blot Analysis

Cell pellets were lysed in a RIPA buffer as described above and loaded to 4% to 12% gradient SDS-PAGE (GenScript) gels by mixing with SDS loading buffer (100 mM pH6.8 Tris-Cl, 4% SDS, 0.2% bromophenol blue, 20% glycerol, 200 mM DTT) after

denaturation at 95°C for 10 minutes. Proteins were transferred using Nitrocellulose membranes (Millipore) and blocked in 5% non-fat milk for 1 hr. Membranes were then incubated with the primary antibodies listed above at 4°C overnight followed by incubation with a secondary antibody at room temperature for 1 hr. West-Q Pico Dura ECL Solution (GenDepot) was used for developing, and the antigen–antibody complexes were detected by the ChemiDoc Imaging system (Bio-Rad).

2.5.9. Loci Specific Methylation Analysis

Amplicon sequencing analysis was performed as described previously. Briefly, MethPrimer[104] was used to design primers against the regions of interest. Purified genomic DNA was treated with sodium bisulfite using an EZ DNA Methylation-Lightning Kit (Zymo Research). The regions of interested were amplified by PCR using a PyroMark PCK kit (Qiagen). The amplified regions were purified by an MinElute PCR purification kit (Qiagen). The sequencing libraries were generated by Nextera XT DNA library prep kit (Illumina). Libraries were sequenced on Illumina NextSeq 500 system. The amplicon sequences were aligned to mm10 genome using BSMAP and then the multi-mapped reads were removed. MOABS[105] was used for DNA methylation calling. We also removed CpGs sites with coverages less than 10.

2.5.10. Flow Cytometry Analysis

Cells were washed with PBS and resuspended in a FACS buffer (PBS, 0.5% BSA, 2 mM EDTA). Antibodies were used as described above and incubated with cells for 30 mins on ice. Cells were then washed with FACS buffer twice and then analyzed by an LSR II cell cytometer (BD Bioscience). Data was analyzed by the FlowJo software (BD).

2.5.11. Single Cell RNA-seq Library Preparation and Data Analysis

EBs collected at Day 5 differentiation procedure were digested into single cells by 0.5% Trypsin (GenDepot). Single cell Gel Bead-In Emulsions (GEM) was generated and barcoded in a Chromium Controller (10x Genomics). Single-cell RNA-seq libraries were prepared using the Chromium Single-Cell 3' Reagent V2 Kit (10x Genomics) according to the manufacturer's instructions. Agilent Bioanalyzer 2100 (Agilent) was used to examine the quality of libraries. Libraries were sequenced using the Illumina NextSeq 500 system.

Cellranger (v3.0.2) was used to demultiplex the raw sequencing data, generate fastq files and align the files to mm10 genome. The UMI counting and barcode decoding for each cell were achieved by using Cellranger. The raw reads count matrix with row as genes and column as cells were input to Seurat (v2.3) to perform quality control and the downstream analysis. Cells with more than 5% mitochondrion reads and detected less than 100 genes were discarded. Genes detected in less than 50 cells were also removed. The default settings were used for UMAP clustering and cluster identification. Monocle, an unsupervised algorithm, was used to order single-cell expression profiles by a quantitative measurement of progress through a biological process (pseudotime trajectory). The default settings were applied on both the control and Tet-TKO datasets. We used scVelo python package to perform single cell velocity analysis. Specifically, we calculated the ratio between pre-mRNA and mRNA for each gene and each cell. The ratio between pre-mRNA and mRNA abundance can be used to infer the future gene expression dynamics in cells.

2.5.12. CMS-IP Library Construction and Data Analysis

CMS-IP-seq was performed to analyze genome-wide 5hmC distribution as previously described. Purified genomic DNA was sheared into 300~500 bp fragments using Covaris (ME220). DNA fragments were treated with sodium bisulfite using an EZ DNA Methylation-Lightning Kit (Zymo Research) to convert 5hmC to cytosine-5-methylsulfonate (CMS). An anti-CMS antibody and protein A/G dynabeads (Thermo Fisher Scientific) were used to pull-down the CMS-containing DNA fragments. Enriched DNA fragments were purified and subsequently used for library preparation with a Pico Methyl-Seq Library Prep Kit (Zymo Research). Libraries were then sequenced using Illumina NextSeq 500.

For CMS-IP–seq analysis, we first aligned raw reads to mm10 using BSMAP [39]. Only uniquely mapped reads were retained for the downstream analysis. A CMS count table was tabulated for the 5hmC-enriched regions detected by MACS2[106] . We used the CMS-IP software to detect DHMRs between different differentiation days (https://github.com/lijinbio/cmsip). CMS-IP computes false discovery rate (FDR)–adjusted q values of the P values in G test using the Benjamini and Hochberg procedures. To reduce the loss of statistical power caused by the FDR adjustment, independent filtering is applied to rule out low-count regions, where the filtering criterion measures the average normalized CMS counts across sample replicates. DHMRs were identified by a $q < 0.05$. To facilitate the visualization of 5hmC signals, bigWig files for read coverage were generated from the aligned BAM files and visualized in the UCSC genome browser.

2.5.13. Accession Numbers and Data Availability

The sequencing datasets were deposited into the NCBI BioProject under the accession numbers GSE168061.

3. TET INACTIVATION DISRUPTS YY1 BINDING AND LONG-RANGE CHROMATIN INTERACTIONS DURING EMBRYONIC HEART DEVELOPMENT[39]

3.1. Summary

Tet-mediated DNA demethylation plays an important role in regulating embryonic development. However, little is known about their functions in mammalian heart development. In this part, we first analyzed DNA methylation and hydroxymethylation dynamics during early cardiac development in both human and mice. By using cardiac development specific knock-out mouse model, we found that deletion of Tet2 and Tet3 in mice (Tet2/3-DKO) lead to ventricular non-compaction cardiomyopathy (NCC) with embryonic lethality. Single-cell RNA-seq analyses revealed a reduction in cardiomyocyte numbers and transcriptional reprogramming in cardiac tissues upon Tet2/3 depletion. Impaired DNA demethylation and reduced chromatin accessibility in Tet2/3-DKO mice further compromised Ying-Yang1 (YY1) binding to its genomic targets, and perturbed high-order chromatin organization at key genes involved in heart development. Our studies provide evidence of the physiological role of Tet in regulating DNA methylation dynamics and chromatin organization during heart development.

3.2. Introduction

Cardiac development is a complex biological process, which is precisely regulated. Epigenetic factors, such as DNA modifying enzymes, play indispensable roles in choreographing this exquisitely coordinated process by directly participating in the programming of cardiac transcriptional networks, thereby exerting control over gene

expression to orchestrate early cardiac development[44,76,107]. Aberrant epigenetic modifications arising from genetic alterations in these key enzymes and / or environmental risk factors (such as folate deficiency) may cause developmental defects in heart and even lead to embryonic lethality in mice and human cardiomyopathies[108-110]. A deeper understanding on epigenetic regulatory mechanisms that modulate cardiac gene expression is crucial for deciphering the molecular etiology of congenital heart defects.

As introduced before, Tet mediated DNA modification pathway plays an important role in regulating DNA methylation and demethylation homeostasis during development. Tet-catalyzed DNA hydroxymethylation is closely associated with enhancer activity and chromatin accessibility during cellular differentiation and embryonic development[40,103,111].

In this study, we systematically evaluated the DNA methylation and hydroxymethylation dynamics during early cardiac development in human and mice. We then generated a cardiac specific Tet-deficient mice model to investigate the function of Tet-mediated DNA methylation and demethylation homeostasis during early cardiac development. With these physiologically relevant in vivo models, we unveiled previously unrecognized functions of Tet-mediated DNA hydroxymethylation in regulating chromatin accessibility to facilitate the genomic recruitment of key transcription factor, YY1 (Ying-Yang 1), as well as long-range chromatin contacts at cardiac-specific genes during embryonic heart development.

3.3. Results

3.3.1. 5mC and 5hmC Undergo Dynamic Changes During Embryonic Heart Development in Both Human and Mice.

To evaluate DNA methylation dynamics during mammalian heart development, we performed whole-genome bisulfite sequencing (WGBS; for 5mC profiling) and CMS-IP-seq (for 5hmC profiling) in both human and mouse embryonic heart tissues. For human heart tissues, we analyzed DNA methylation and hydroxymethylation at the Carnegie stage (CS) 13 and 14, which are equivalent to embryonic day 9.5 (E9.5) to E10.5 of the murine heart development stages[112]. For mouse hearts, we systematically analyzed DNA methylation dynamics using WGBS data that are available in ENCODE[113] at different embryonic heart development stages (E11.5, 12.5, 13.5, 14.5, 15.5, 16.5, P0). We also performed CMS-IP-seq in E12.5 mouse embryonic cardiac tissues to correlate 5mC with 5hmC levels at this stage.

We first comprehensively analyzed the DNA methylation dynamics in mouse hearts since ENCODE data covered most of the key cardiac development stages. Although the global DNA methylation levels remained largely stable during embryonic development by ranging from 0.736 to 0.755 across all the stages (Figure 3.1A), we were still able to locate 21,467 differentially methylated regions (DMRs, defined as >20% methylation change, FDR $\leqslant$ 0.05) that covered 105,710 CpG sites from E11.5 to P0 (Figure 3.1B), suggesting active dynamic changes in focal rather than global DNA methylation during cardiac development. GREAT analysis on all the identified DMRs further revealed that these regions are enriched at cis-regulatory region of genes that are essential for embryonic development and heart function, such as Bmp10 and Tnnt2 (Figure 3.1C-D). Among all

identified DMRs between adjacent developmental stages, approximately 66% of DMRs (14,155 out of 21,476 DMRs) exhibited a reduction in DNA methylation when progressing to the next developmental stage (defined as hypoDMRs; Figure 3.1B). Notably, more than 80% DMRs were classified as hypoDMRs during the E12.5-to-E13.5, or E16.5-to-P0 transitions (Figure 3.1B), implying that DNA hydroxymethylation (with ultimate demethylation) probably is a dominant phenomenon at these developmental stages.

We next measured the global changes of 5hmC by a well-established dot-blot assay in murine heart tissues collected at different development stages (E12.5, 14.5, 16.5, 18.5 and P0). We observed a gradual increase of global 5hmC during heart development (Figure 3.1E), suggesting that 5hmC might be involved in regulating DNA methylation dynamics in this development process. To further study the correlation between 5mC and 5hmC at loci-specific levels, we isolated heart tissue from E12.5 mouse embryo and performed CMS-IP for genome-wide 5hmC analysis. We observed strong 5hmC enrichment at identified DMRs during heart development, with higher enrichment of 5hmC at hypoDMRs than at hyperDMRs ('hyperDMR' defined as DMRs with significant gain of 5mC signals when transition into the next developmental stage; Figure 3.1F), suggesting that 5hmC is involved in regulating DNA methylation dynamics at these regions during embryonic heart development. We further analyzed 5hmC enrichment in human embryonic heart tissue at CS13 and 14. We also observed a pronounced increase of 5hmC at CS13-to-CS14 transition during human embryonic heart development (Figure 3.1G). Congruently, these epigenomic analysis results validated that focal 5mC changes likely

arising from altered DNA hydroxymethylation are closely associated with embryonic heart development in both mice and human.

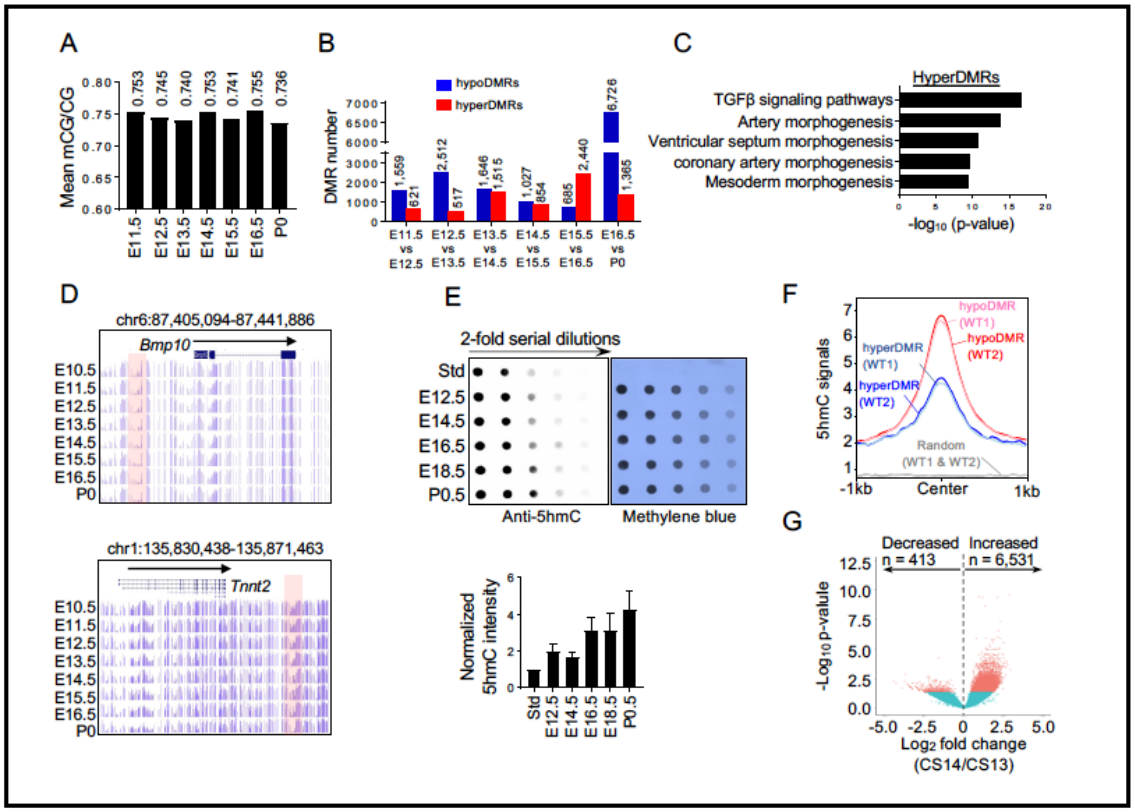

Figure 3.1.Dynamic Changes of DNA Methylation and Hydroxymethylation During Embryonic Heart Development in Mmice and Human.

A) Quantification of global average DNA methylation levels (quantified as mean mCG/CG) during mouse heart development (E11.5 to P0) based on WGBS data from ENCODE;B) Numbers of differentially methylated regions (DMRs) that show increased (defined as hyperDMRs) or reduced DNA methylation (hypoDMRs) when mouse embryonic hearts progress into the next developmental stage; C) GREAT analysis on hyperDMRs shown in Figure 3.1B; D) Representative genome browser views illustrating the focal DNA methylation dynamics at genomic regions surrounding Bmp10 and Tnnt2 during mouse embryonic cardiac development (E10.5 to P0); E) (Upper) Global 5hmC

levels in mouse heart tissues collected at five embryonic stages (E12.5, 14.5, 16.5, 18.5, and P0) measured by the dot-blot assay. Methylene blue is used as loading control for total DNA input. (Lower) Quantification of dot-blot assay. Data were shown as mean ± S.D; n = 3 independent experiments; F) 5hmC enrichment signals of WT E12.5 heart tissues within hypoDMRs (two repeated samples; red and pink) and hyperDMRs (two repeated samples; blue and cyan) identified from Figure 1B. Random genomic regions (grey) were used as control; G) The Volcano plot of differentially enriched 5hmC regions (DHMRs; CS13 vs CS14) in human embryonic hearts.

3.3.2. Cardiac-Specific Tet Depletion Led to Ventricular Non-Compaction Cardiomyopathy (NCC) in Mice.

To further unravel the function of Tet protein mediated DNA methylation oxidation during embryonic heart development, we set out to generate a cardiac tissue-specific Tet-deficient mouse model. Given that Tet1 knockout mice had no obvious cardiac phenotypes[114], we focused on dissecting the role of Tet2 and Tet3 in this study. Meanwhile, straight knockout of Tet3 resulted in embryonic lethality to prevent systematic studies on cardiac development[68]. To avoid these caveats, we decided to cross mice bearing a conditional Tet3flox allele with the cardiomyocytes (CM) progenitor driver line, Nkx2.5-Cre[115], to yield cardiac-specific deletion of Tet3. Next, we crossed Tet3flox; Nkx2.5-Cre mice with Tet2KO mice to generate double lesions of Tet2 and Tet3 genes in CM progenitors (named as Tet2/3-DKO). While the heterozygous mice are viable and showed no obvious cardiac phenotypes, homozygous mice turned out to be embryonically lethal (Figure 3.2A). Therefore, we collected embryos at development stages starting from E12.5 for further analyses.

Timed mating experiments by crossing Tet2+/-Tet3flox; Nkx2.5-Cre male mice with Tet2-/-Tet3flox female mice further revealed a decreased Tet2/3-DKO embryo numbers starting from E15.5 (Figure 3.2A). WT-Nkx2.5-Cre (+/-) embryos as the control. After histological analysis on embryonic hearts collected at E12.5, 13.5, 14.5 and 15.5 stages, we found that Tet2/3-DKO hearts displayed noticeable developmental defects, including ventricular septal defect (VSD) and double outlet right ventricle (DORV) (data not shown). Tet2/3-DKO hearts showed abnormal ventricular chamber development starting from E13.5 (Figure 3.2B). Specifically, Tet2/3-deficient hearts showed a ventricular non-compaction cardiomyopathy (NCC) phenotype, as evidenced by significantly reduced ventricular wall thickness and increased trabecular areas compared to the WT group at the same development stage (Figure 3.2B). Next, we performed real-time quantitative PCR to examine the expression of signature genes (Nppa and Hey2; Figure 3.2C) that are indicative of ventricular NCC in Tet2/3-DKO hearts. We detected a substantial decrease of Nppa and Hey2 expression in Tet2/3-DKO, thus validating the NCC phenotype at molecular level. Because ventricular NCC has been shown to arise from CM proliferation defects and / or increased cellular apoptosis[116,117], we further performed immunofluorescent (IF) staining with cellular proliferation and apoptotic markers in WT and Tet2/3-DKO E12.5 hearts. We observed a significant reduction in the staining for a cellular proliferation marker, Ki67 (Figure 3.2D). Regarding cleaved caspase-3 as apoptotic marker, we failed to detect meaningful signals in E12.5 WT and Tet2/3-DKO heart tissue (Figure3.2E). Together, these results clearly established the critical roles of Tet2 and Tet3 in mediating normal ventricular chamber development.

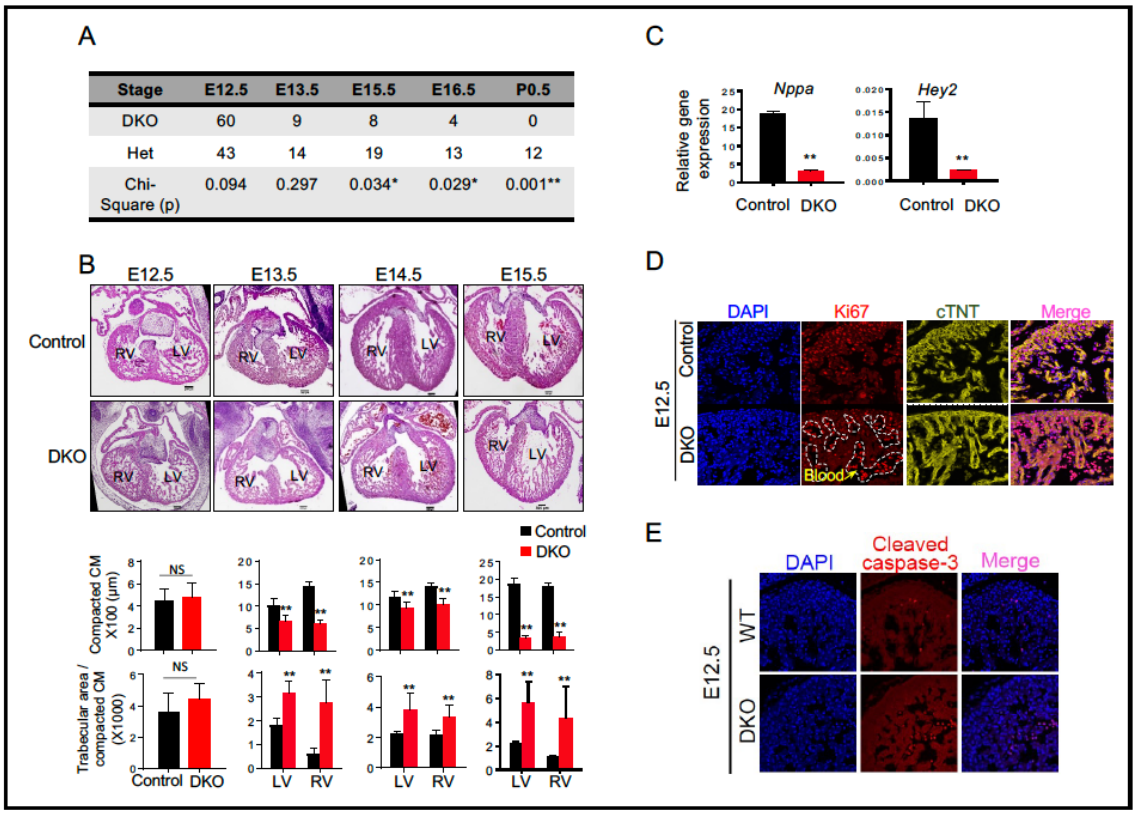

Figure 3.2.Cardiac-Specific Deletion of Tet2 and Tet3 Resulted in Developmental Defects in the Ventricular Chamber.

A) Survival rates of cardiac-specific Tet2 and Tet3 deficient mice at different developmental stages; B) Representative H&E staining images of embryonic heart tissues (4X) collected at E12.5, 13.5, 14.5, and 15.5 stages from Control and Tet2/3-DKO mice. Quantifications were performed by using the ImageJ software. Data were shown as mean ± S.D; n = 36 sections from 3 independent experiments. ** $p < 0.01$ compared to Control; C) Real-time qPCR to quantify the expression of Nppa and Hey2 in Control and Tet2/3-DKO embryonic heart tissues collected at E12.5. Data were shown as mean ± S.D; n = 3 independent experiments. ** $p < 0.01$; D) Representative fluorescent imaging of embryonic heart tissues collected from E12.5 of Control (top) or Tet2/3-DKO mice

(bottom). DAPI (blue) was used for nuclear staining; Ki67 (red) was used as a proliferation marker; and cTnT (yellow) was used for the staining marker for cardiomyocytes. Cells demarcated within the white dashed lines are blood cells; E) Representative fluorescent imaging of embryonic heart tissues collected from E12.5 of WT (top) or Tet2/3-DKO mice (bottom). DAPI (blue) was used for nuclear staining; Cleaved Caspase3 (red) was used as the apoptosis marker.

3.3.3. Tet2 and Tet3 Regulates Cardiac-Specific Transcription During Heart Development

To mechanistically address how Tet2 and Tet3 regulates embryonic heart development, we set out to perform transcriptomic profiling with RNA-seq in embryonic heart tissues collected from E12.5 and E15.5. We identified a total of 2,096 differentially expressed gene (DEGs), with 1,263 down-regulated and 833 up-regulated genes, in Tet2/3 DKO samples collected at E12.5 embryos (Figure 3.3A). In parallel, we identified 440 and 374 up and down-regulated genes in E15.5 embryonic hearts (Figure 3.3A). Gene ontology (GO) analysis on these DEGs revealed the involvement of key signaling pathways (e.g., Notch related signaling) that are known to be critical for ventricular chamber development (Figure 3.3B)[116,117]. These unbiased transcriptomic and bioinformatic analysis data provided further evidence to support a critical role of Tet2/3 in embryonic heart development.

To further examine the function of these Tet2/3-regulated DEGs identified from RNA-seq analysis, we sorted DEGs implicated in cardiac development based on GO analysis (Figure 3.3C). Then we analyzed the expression level of these genes using ENCODE RNA-seq data collected from mouse embryonic hearts at different development stages

(E10.5 to P0; Figure 3.3C). Interestingly, we found that the expression levels of these genes underwent gradual changes during heart development, suggesting that these Tet2/3-regulated genes are tightly and temporally controlled at different embryonic stages. Deletion of Tet proteins might disrupt the precise transcription regulation of these key genes to impair cardiac development.

The above RNA-seq analyses were performed in bulky embryonic heart tissues collected at E12.5 and E15.5 that contain multiple cell types, including myocardium, epicardium, endocardium, fibroblasts and other non-CM cells (e.g., hematopoietic cells). To avoid potential biased results due to changes in cell types upon Tet2/3 knockout, we decided to carry out single-cell RNA-seq (scRNA-seq) analyses on WT and Tet2/3-DKO E12.5 and E15.5 hearts. Consistent with our histological analysis data (Figure 3.2C), we observed minor difference in the cell types in E12.5 hearts between WT and Tet2/3-DKO embryos. However, we noticed approximately 5-fold reduction in the myocardium population (from 32.9% to 6.8%) in the Tet2/3-DKO group at E15.5 (Figure 3.3D). Furthermore, consistent with our real-time qPCR analysis and bulk RNA-seq data described above (Figure3.2D and Figure 3.3A), key genes involved in ventricular NCC, such as Nppa and Hey2, are significantly down-regulated in Tet2/3-DKO hearts collected at E15.5 when compared with the WT group (Figure 3.3E). Furthermore, CM signature genes, such as Tbx20, Ttn and Gja1, were significantly down-regulated in Tet2/3-DKO E15.5 hearts (Figure 3.3F). Together, these data confirmed that Tet2 and Tet3 are essential for maintaining myocardium development.

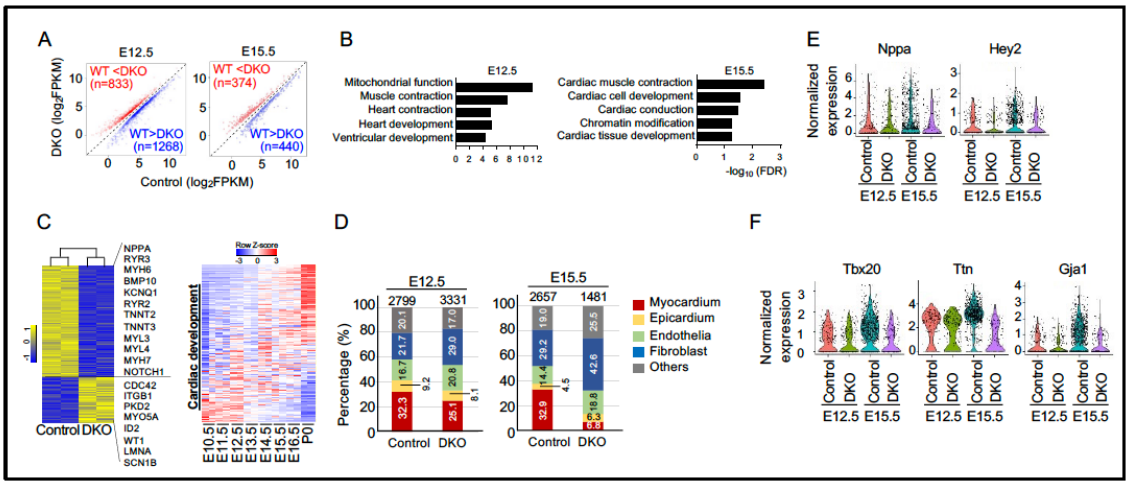

Figure 3.3.Transcriptomic Analyses on Control and Tet2/3-DKO Embryonic Heart Tissues.

A) Scatter plot of the RNA-seq expression data to identify differentially expressed genes (DEGs) in embryonic heart tissues between the Control and Tet2/3-DKO groups at the E12.5 (left) or E15.5 (right) developmental stages. DEGs were defined as q-value <= 0.05. Red and blue dots stand for up- and down-regulated genes, respectively, in the Tet2/3-DKO group when compared to WT; B) GSEA analysis of DEGs identified between Control and Tet2/3-DKO embryonic heart tissues collected at E12.5 and E15.5; C) (Left) Heatmap presentation of the same group of cardiac development-associated DEGs in Control and Tet2/3-DKO heart tissue collected at E12.5. (Right) Heatmap presentation of expression data for cardiac development-associated DEGs (identified using GSEA analysis between E12.5 Control and Tet2/3-DKO groups) in embryonic heart tissues collected at different developmental stages (E10.5 to P0). RNA-seq data were obtained from ENCODE; D) Percentages of cell types in E12.5 and E15.5 heart tissues collected from Control and Tet2/3-DKO mice using single-cell RNA-seq (scRNA-seq) analysis.

Numbers listed above each bar represent the total analyzed cell numbers; E) Violin plot showing the distribution of normalized expression levels of selected DEGs at E12.5 and E15.5 heart tissues collected from the Control and Tet2/3-DKO groups. Each dot represents the expression levels of corresponding genes in single cells.

3.3.4. Tet2/3 Depletion Causes a Global Decrease of 5hmC with Global 5mC Levels Relatively Unaltered in Embryonic Hearts

The dot-blot assay confirmed a substantial decrease of 5hmC in Tet2/3-DKO heart tissues when compared to the control group (Figure 3.4A). To further delineate the function of 5hmC in regulating cardiac-specific gene expression, we performed CMS-IP-seq analysis to profile genome-wide 5hmC in E12.5 heart tissues. Consistent with dot-blot result, we observed a global decrease of 5hmC in Tet2/3-DKO samples (Figure 3.4B). Among 9,559 identified differentially hydroxymethylated regions (DHMRs) between the WT and Tet2/3-DKO groups, we identified 8,846 genomic regions with reduced hydroxymethylation (defined as hypoDHMRs) and only 713 regions with increased DNA hydroxymethylation (hyperDHMRs) in Tet2/3-KO heart tissues (Figure 3.4B). GREAT analysis unveiled that hypoDHMRs were primarily enriched at distal-regulatory regions of genes that are important for heart development and function (Figure 3.4C). Integrative analysis was performed among our 5hmC mapping data, the histone enrichment datasets from ENCODE and DNA methylation datasets from ENCODE. We found that DNA methylation level at these Tet2/3-regulated hypoDHMRs were relatively low and these regions were highly enriched with H3K4me1 and H3K27Ac, which are usually marked at enhancers (Figure 3.4D). By comparison, the average DNA methylation of hyperDHMRs were relatively high, accompanied by moderate enrichment of H3K36me3 but no other

histone marks (Figure 3.4D). Congruently, these data suggested that Tet2/3 loss associated reduction of 5hmC is more associated with distal regulatory regions at cardiac development-associated genes.

Next, we aimed to address how Tet/5hmC loss alters the transcription of cardiac-specific genes during cardiac development. We measured DNA methylation in Tet2/3-DKO heart tissues collected at E12.5 using whole genome-wide bisulfite sequencing (WGBS) analysis. In parallel, we compared our WGBS data with ENCODE WGBS data collected with E12.5 WT heart tissues. We did not detect a significant change in the average DNA methylation level (mCG/CG) between the two groups (Figure3.4E). We identified only 388 hyperDMRs and 19 hypoDMRs in E12.5 Tet2/3-DKO heart tissues when compared to the control samples (Figure 4F), suggesting that Tet2/3 deletion only exerted a minor influence on the overall DNA methylation levels. These results indicate Tet loss induced focal DNA methylation changes might not be the main reason responsible for the observed ventricular NCC phenotype.

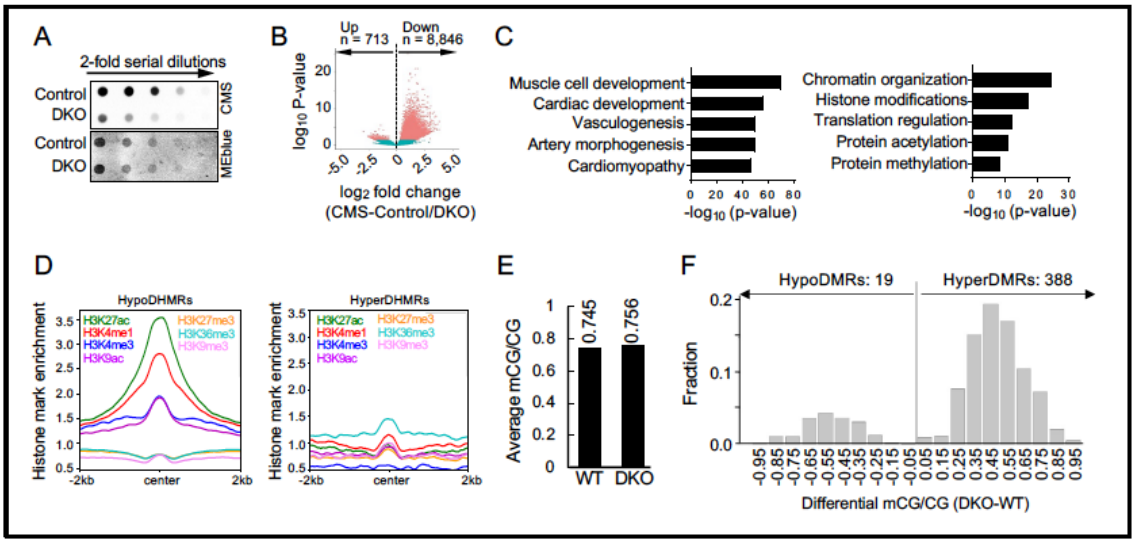

Figure 3.4.Tet2/3 Deletion in Embryonic Heart Resulted Impaired 5hmC but not 5mC.
A) Measurement of the global 5hmC levels in Control and Tet2/3-DKO heart tissues collected at E12.5 by using the dot-blot assay. Methylene blue (MEblue; bottom) staining was used to visualize the total DNA input; B) Volcano plot illustrating the differentially enriched 5hmC regions (DHMRs) in E12.5 heart tissues between the Control and Tet2/3-DKO groups (p value <=0.05); C) GREAT analysis on DHMRs illustrated in Figure 3.4B; D) Normalized enrichment of the indicated histone modifications within identified hypoDHMRs (left) and hyperDHMRs (right); E) Average DNA methylation levels in WT and Tet2/3-DKO murine heart tissues collected at E12.5; F) Histogram showing the difference of mCG/CG ratio at each CpG in the WT and Tet2/3-DKO groups.

3.3.5. 5hmC Loss Reduces YY1 Binding to Its Genomic Targets by Modulating Chromatin Accessibility in Embryonic Cardiac Tissues

Together with other groups, we reported the connection between Tet protein mediated 5hmC in regulating chromatin accessibility[40,103,111]. This led us to further examine the

genome-wide changes of chromatin accessibility by using ATAC-seq in Tet2/3-DKO embryonic heart tissue. Upon Tet2/3 deletion, a total of 2,816 and 960 genomic regions displayed reduced or increased chromatin accessibility, respectively, in mouse embryonic hearts (Figure 3.5A). Interestingly, we observed a strong positive correlation between 5hmC, and chromatin accessible regions as reflected by a Pearson correlation coefficient of 0.85 in both WT and Tet2/3-DKO heart tissues (Figure 3.5B). Subsequently, we selected genomic regions displaying altered 5hmC or chromatin accessibility in the Tet2/3-DKO group and found that 68.4% of selected regions showed simultaneous reduction in 5hmC and chromatin accessibility (Figure 3.5C). GREAT analysis revealed that these overlapping regions were mostly enriched at distal regulatory regions of genes (such as Mly2, Tnnt2, and Ttn) that are involved in regulating Notch signaling and heart development (data not shown). We next examined the relationship between 5hmC/ATAC-low regions and RNA-Seq signals for the related genes and detected a strong positive correlation. For instance, we observed a strong association between altered expression of cardiac development genes (e.g., Nppa, Tnni2, and Bmp10) and reduced 5hmC/ATAC-seq signals within 1kb of transcription start sites (TSS) at corresponding genes (Figure 3.5D-E). Taken together, these data suggest that cardiac specific deletion of Tet protein induced 5hmC loss and caused a reduction in chromatin accessibility to alter cardiac gene expression.

Based on the data described above, we hypothesized that Tet deletion reduces 5hmC and chromatin accessibility, and subsequently affects key TF binding to their genomic targets in embryonic hearts. GREAT analysis on ATAC-seq data from WT and Tet2/3-DKO heart

tissues pointed to YY1 as the top candidate, which showed strong enrichment in genomic regions displaying reduced chromatin accessibility in the Tet2/3-DKO group (Figure 3.5F). YY1, a member of the Gli-Kruppel family of zinc finger protein, is an important transcription factor regulating early heart development[118]. We tested these findings in our Tet-TKO mESC that generated previously. Since Tet-TKO mESC has an undetectable 5hmC level, it provides a clean system to elucidate the impact of 5hmC on TF binding to chromatin. Next, we measured the chromatin association of YY1 in WT and Tet-TKO mESCs. We observed a significant decrease in chromatin associated YY1 in Tet-TKO mESCs compared with parental WT mESCs (Figure 3.5G). To further examine the correlation between 5hmC and YY1, we applied the CUT&RUN (C&R) method[119] to enrich YY1 genomic binding regions in WT and Tet-TKO mESCs. We examined the correlation between YY1 and 5hmC in parental WT mESCs. We observed a strong 5hmC enrichment with concomitant depletion of 5mC signals at YY1 enriched regions (Figure 3.5H). About 20% of YY1-enriched peaks (n = 10,450) overlapped with 5hmC enriched regions (Figure 3.5I). To further examine the impact of Tet/5hmC loss on YY1 genomic distribution, we compared YY1 enriched regions between the WT and Tet-TKO groups and noted a pronounced reduction of YY1 enrichment in Tet-TKO mESCs (Figure 3.5J), particularly at genomic regions that displayed 5hmC reduction upon Tet deletion. These data further confirmed that Tet and 5hmC regulates chromatin accessibility to facilitate the binding of proper TFs (YY1 in this case) to their targets.

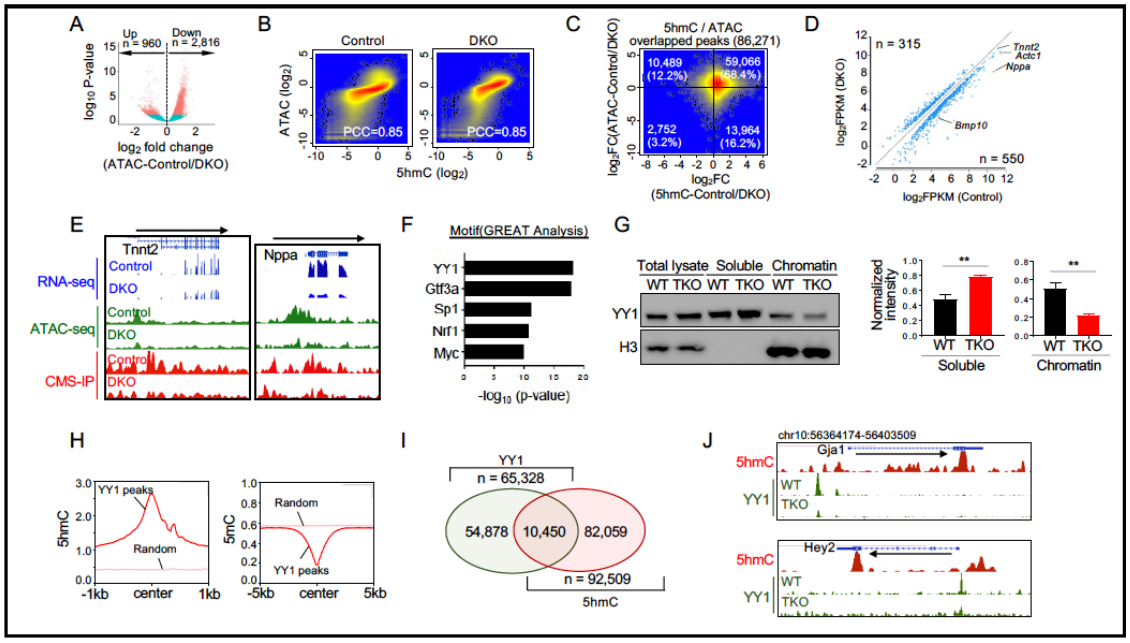

Figure 3.5.Tet2/3 Deletion in Murine Embryonic Hearts Reduced Chromatin Accessibility and Compromised the Binding of YY1 to Chromatin.

A) Volcano plot depicting the differentially ATAC-seq enriched regions between the Control and Tet2/3-DKO groups (p value <=0.05). Data were collected from E12.5 embryonic heart tissues obtained from Control and Tet2/3-DKO mice; B) Scatterplot depicting the correlation between 5hmC and ATAC-seq signals in the Control (left) and Tet2/3-DKO (right) groups. Spearman correlation coefficient was calculated by correlating 5hmC signals with ATAA-seq signals obtained from E12.5 cardiac tissues. The ATAC and 5hmC signals of every 10 kb bin were calculated; C) Scatterplot depicting the distribution of genomic regions with altered 5hmC (x-axis) and ATAC-Seq signals (y-axis) between the Control and Tet2/3-DKO groups; D) Scatterplot presenting the expression of DEGs that showed alterations in 5hmC and ATAC-seq signals within 1 kb of transcription start sites (TSS). Tnnt2, Actc1, Nppa and Bmp10 were among the hits; E) Genome browser view examples of decreased transcription (blue, RNA-seq), chromatin

accessibility (green, ATAC-seq) and 5hmC (red, CMS-IP) enrichment at cardiac-specific gene loci (Tnnt2 and Nppa) in Tet2/3-DKO E12.5 heart tissues compared with that in Control group; F) Motif analysis of genomic regions displaying decreased ATAC seq signals in the Tet2/3 DKO group; G) (Left) Representative western-blot analysis of YY1 binding to chromatin in WT and Tet-TKO mESCs. (Right) Quantification of the intensity of YY1 soluble and chromatin associated fractions in WT and Tet-TKO mESCs. Data were shown as mean ± S.D; n = 3 independent experiments. ** $p < 0.01$ compared to WT; H) Enrichment of 5hmC (left) and 5mC (right) signals within YY1-enriched peaks. Random regions were used as control; G) Venn diagram showing the degree of peak overlaps between YY1- and 5hmC-enriched regions in mESCs. 10,450 overlapping peaks were identified; I) YY1 enrichment in WT (red and green) and Tet-TKO (yellow and blue) mESCs; J) Exemplary genome-browser view of 5hmC enrichment in WT mESCs (red trace) and YY1 enrichment (green) at the Gja1 and Hey2 loci in WT and Tet-TKO mESCs.

3.3.6. 5hmC loss disrupts higher-order chromatin structures in embryonic hearts

YY1 is known to regulate chromatin higher-order structures by controlling promoter-enhancer looping[120]. In addition, together with others, we have reported that 5hmC is enriched at euchromatin[103,111]. Chromatin is known to be spatially categorized into two types of large compartments, A and B, that exhibit either open chromatin domain (A) or closed chromatin domain (B)[121].To examine whether Tet mediated DNA hydroxymethylation is associated with the organization of chromatin higher-order structures, we performed HiChIP experiment in WT and Tet2/3-DKO E12.5 heart tissues using an anti-smc1 antibody[122]. Interestingly, we observed a strong enrichment of 5hmC

in compartment A, but not in compartment B (Figure 3.6A), suggesting that 5hmC tends to mark transcription active regions. Then we compared three-dimensional chromatin interaction patterns between WT and Tet2/3-DKO heart tissues. In general, the compartment organization between WT and Tet2/3-DKO heart showed very similar patterns (R = 0.93; Figure 3.6B). However, we observed 1,424 bins (50 kb resolution) switching from compartment A to B upon Tet2/3 depletion in cardiac tissues (Figure 3.6B). Next, we evaluated the expression levels of genes that fell into the A-to-B compartment switch category. We identified 250 down-regulated genes, with many of them known to be important for heart development, such as Ttn, Cav1, Bmp5 and Actc1 (Figure 3.6C); Go Ontology (GO) analysis showed that these genes are important for maintaining normal heart function or are closely implicated in cardiomyopathies (Figure 3.6D).

Next, we calculated unique paired-ended tags (PETs) using the Fit-HiC pipeline[123] to identify the long-distance contacts in E12.5 heart tissues. We identified 475,630 and 347,816 confident contacts in WT and Tet2/3-DKO heart tissue, respectively (Figure 3.6E). We further moved on to examine the potential impact of altered contacts on gene transcription by counting PETs at DEGs identified between the WT and Tet2/3-DKO groups. We noticed that 60% (n=979) DEGs with mapped PETs exhibited reduced long-distance contacts in the Tet2/3-DKO group (Figure3.6F). For example, at Tbx20 and Hey2 loci (two genes displayed significant downregulation upon Tet deletion in scRNA-seq analysis), we detected a pronounced decrease of promoter-enhancer interactions in the Tet2/3-DKO group when compared with those in the control group (Figure 3.6F). Among genomic regions that displayed decreased promoter-enhancer interactions, we also noticed

a concomitant reduction in 5hmC and ATAC-seq signal enrichment in the Tet2/3-DKO group (Figure 3.6G). These data established that Tet proteins contribute to the formation of long distance promote-enhancer looping as an effective way to regulate gene transcription.

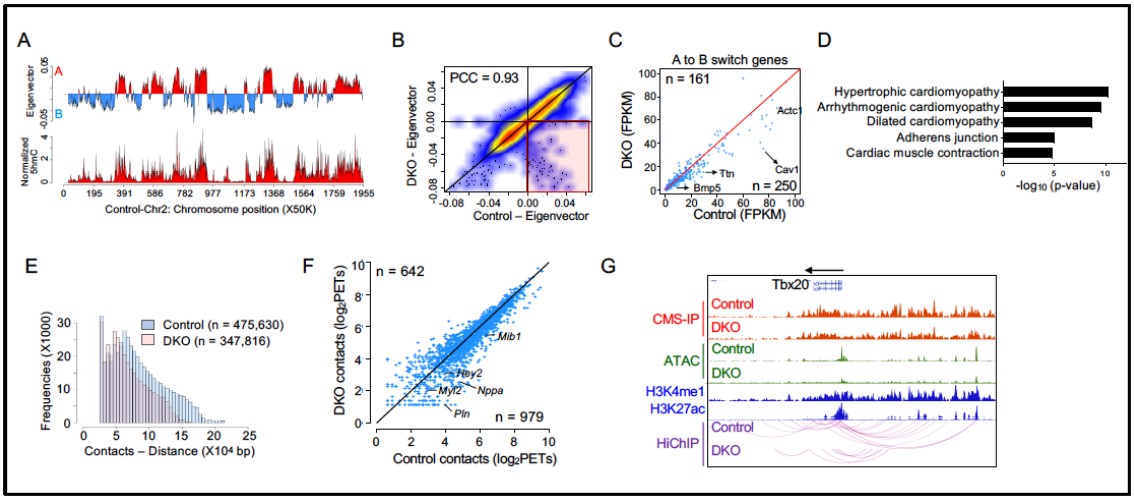

Figure 3.6.Tet-Mediated DNA Hydroxymethylation Is Associated with Chromatin Higher-Order Structures.

A) Representative histograms of compartment A and B distribution (top) calculated by Eigenvector using HiChIP and CMS-IP-seq (for 5hmC) data over chromosome 2 in E12.5 Control embryonic heart tissues; B) Scatterplot showing the distribution of Eigenvector calculated in Control and Tet2/3-DKO embryonic heart tissues at E12.5. Pearson correlation coefficient was calculated; C) Expression levels of genes that underwent A-to-B compartment switch in Control and Tet2/3-DKO embryonic heart tissues (E12.5); D) GSEA analysis of genes located in genomic regions that showed A-to-B compartment switch in Tet2/3-DKO embryonic heart tissues (E12.5); E) Histograms of PETs (distance and numbers) from HiChIP data obtained from Control and Tet2/3-DKO heart tissues collected at E12.5; F) Scatterplot showing the normalized numbers of PETs that

overlapped with DEGs (regions ranging from TSS to TTS) in Control and Tet2/3-DKO heart tissues collected at E12.5; G) WashU Epigenome browser views of 5hmC, ATAC-seq, PETs (HiChIP), H3K4me1 (ENCODE) and H3K27ac (ENCODE) at the Tbx20 locus in Control and Tet2/3-DKO heart tissues collected at E12.5.

3.4. Discussion

DNA methylation is one of the most important epigenetic events controlling heart development and cardiac functions[44,107]. But how the change of DNA methylation results in abnormal heart development is still unknown. In this study, we have systematically analyzed DNA methylation and hydroxymethylation dynamics during embryonic heart development in both human and mouse. Our integrative studies have unveiled dynamic focal DNA methylation changes that are closely associated with 5hmC enrichment at genes essential for cardiac development.

TET protein family is one of the major regulators controlling DNA methylation oxidation. The current study is exclusively focused on studying Tet2 and Tet3 given their relatively higher expression in embryonic cardiac tissues. Cardiac specific deletion of Tet2 and Tet3 resulted in ventricular non-compaction cardiomyopathy in mice, clearly attesting to the indispensable roles of these two epigenetic modifiers in normal heart development. At the molecular level, deletion of Tet2/3 prominently impaired DNA hydroxymethylation in heart tissue and altered the transcription of cardiac development associated pathways, such as Notch pathways. Mechanistically, we propose that changes in chromatin accessibility, likely attributed to compromised DNA hydroxymethylation, sabotage the binding of key cardiac development-associated TFs, as exemplified by YY1, to their targets across the genome. YY1 is a key mediator of embryonic heart development by facilitating GATA4

mediated transcriptional activation and promoting cardiac progenitor cell commitment[118,124]. Interestingly, cardiac specific YY1 knockout mice (YY1f/f-Nkx2.5Cre) displayed very similar phenotypes as Tet2/3-DKO mice generated in the current study, with both in vivo models showing reduced embryo survival at E13.5 and decreased proliferation of cardiomyocytes (with similar non-compaction cardiomyopathy manifestations)[125]. These findings strongly suggest that Tet proteins and YY1 might converge to regulate the similar transcription regulatory pathways during embryonic cardiac development.

In addition to Tet2 and Tet3, we also observed dynamic changes of Tet1 expression during heart development in both human and rodents. We cannot rule out the possibility that Tet1 might be also an important contributor to embryonic heart development. Nonetheless, no heart-related phenotypes have thus far been reported upon Tet1 deletion in transgenic mice, probably due to the redundant functions of Tet homologs (Tet2 and Tet3). Further studies on a cardiac-specific Tet triple knockout mouse model might address the additional function of Tet1 during embryonic heart development.

3.5. Materials and Methods

3.5.1. Human Samples

Human embryos collecting protocol was approved by the Ethical Internal Review Board of the Xinhua Hospital, Shanghai, China. Human embryos were collected from pregnant mothers who performed clinical drug abortion at the Department of Obstetrics and Gynecology in Xinhua Hospital. Consent forms were signed. The embryonic stages of the embryos were measured by using a standard protocol reported previously[126].

3.5.2. Animal Models

Animal studies were approved by the Institutional Animal Care Use Committee (IACUC) of the Institute of Biosciences and Technology, Texas A&M University. Most mouse strains bear a C57BL/6 genetic background unless otherwise noted. Tet2-/-[127], Tet3f/f[70] and Nkx2.5-Cre (The Jackson Laboratory 024637)[128] mouse strains were reported previously. Timed pregnancies were applied and the day on which a plug was found was defined as E0.5. Mice tails were cut and boiled in 50 mM NaOH for 1h and then neutralized in 10 mM Tris-HCL at pH7.4. PCR was carried out using the EmeraldAmp GT PCR Master Mix (TaKaRa) according to the manual. Genotyping primers were listed in Table 3.1.

Table 3.1.Genotyping Primers of Tet, Tet3 and Nkx2.5Cre Mice

Gene target	Direction	Sequences (5'->3')
Tet2 WT	Forward	GCCCAAGAAAGCCAAGACCAAGAA
	Reverse	AAGGAGGGGACTTTTACCTCTCAGAGCAA
Tet2 Deletion	Forward	CGTGTGTCATGCATTTCTGTGTGC
	Reverse	TCGACGGTATCGATAAGCTTCGGATC
Tet3 WT	Forward	GGATGTGAGCTAGTTCTCCTAACTTGAGAGG
	Reverse	CCCTGTCTACTCTATTCTTGTGTCAGGAGG
Nkx2.5Cre	Forward	TTACGGCGCTAAGGATGACT
	Common	GAGCCTGGTAGGGAAAGAGC
	Reverse	GTGTGGAATCCGTCGAAAGT

3.5.3. Antibodies

For IHC: Tet1 antibody was kindly provided by Dr. Leonhardt Heinrich (1:100)[129]. Anti-Tet2 (Abcam ab124297, 1:100), anti-5mC (Millipore MABE146, 1:1,000), anti-5hmC (Active Motif 39769, 1: 40,000), anti-5fC (Active Motif 61223, 1:2,000), and anti-5caC (Active Motif 61225, 1:1,000) were purchased from commercial sources.

For IF: anti-Ki67 (Abcam ab16667, 1:100), anti-cleaved caspase-3 (Cell Signaling Technologies 9661s, 1:50); Alexa Fluor 568 goat anti-rabbit (Thermo Fisher Scientific A-11011, 1:1,000), Alexa Fluor 647 goat anti-mouse (Thermo Fisher Scientific A-21235, 1:1,000)

For CUT&RUN experiments: anti-YY1 (Santa Cruz sc-7341, 1:100), anti-H3K27ac (Abcam ab4729, 1:100), Rabbit anti-mouse (Abcam ab6728, 1:100), anti-cTNT (Thermal Fisher Scientific 13-11, 1:100) were purchased.

For HiChIP: anti-Smc1 (A300-055A, Bethyl Laboratories, Inc. 2 μg / reaction) was used.

For Western Blotting: anti-YY1 (Santa Cruz sc-7341, 1:1000), and anti-H3 (Abcam ab1791, 1:2,000) antibodies were used.

3.5.4. Histological Analysis

All mouse embryos were dissected in phosphates buffered saline (PBS). Embryos pictures were taken using a Nikon SMZ800N dissecting microscopy. For histological characterization, all embryos were fixed overnight in 4% PFA, then dehydrated with graded ethanol and embedded in paraffin. Sections were cut at the thickness of 7 μm. Slides were dried at 37˚C overnight and then stained with hematoxylin-eosin (H&E) as

previously described[130,131]. Stained sections were imaged using a Nikon Eclipse Ci microscopy.

3.5.5. Quantification of the Compact Myocardium Thickness and the Trabecular Area

ImageJ was used for the measurement and quantification of histological data. The quantification method was described previously[116]. In briefly, the ventricle was divided into the apex and the basal regions, and several measurements of the compact myocardium thickness were taken for each region. The average thickness was calculated. The trabecular area was measured by dividing the surface occupied by the trabeculae in the ventricle. The ratio of the trabecular area and compact myocardium thickness was used as an indicator for the size of the trabecular mesh.

3.5.6. Immunohistochemistry (IHC) and Immunofluorescence (IF)

IHC and IF were performed as previously described[130,131]. Briefly, tissue sections were dewaxed in xylene twice for 5 min each and rehydrated in a graded series of ethanol (100% to 70%). The antigens were retrieved by boiling sections in 10mM citrate buffer (Vector Laboratories) for 20 min. For DNA modifications staining, sections were treated with 2N HCl for 30 min to expose the epitopes and then neutralized in 100 mM Tris-HCl (pH 8.5) for 10min. IHC was using the ImmPRES HRP Reagent Kit (Vector Laboratories) to perform the blocking and antibody incubation, and then developed by using the DAB peroxidase substrate kit (Vector Laboratories). IHC stained sections were imaged by a Nikon Eclipse Ci microscopy. For IF, 10% normal goat serum (Thermo Fisher Scientific) was used to block the unspecific antigens. After primary and 2nd antibodies incubation, 0.5 μg/ml DAPI (Thermo Fischer Scientific, D1306) was used to co-stain the nuclei. IF stained sections were imaged using a Nikon A1 confocal microscope.

3.5.7. Nuclear Fractionation and Western Blot

Nuclear fractionation in mESCs was performed as described previously[132]. Briefly, 10 million cells were washed by PBS and pelleted at 200 g for 2 min. 200 µl buffer A (10 mM HEPES, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10 % Glycerol, 0.1% Triton X-100, 1 mM DTT, and protease inhibitor cocktail) was added to the cell pellets and incubated on ice for 8 min to remove the cytoplasm. After centrifugation at 1,300 x g, 4˚C, for 5 min, 100 µl Buffer N (15 mM Tris-HCl [pH 7.5], 200 mM NaCl, 60 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.3% NP-40, and protease inhibitor cocktail) was added to the nuclear pellets and incubated on ice for 30 min to lyse. After centrifugation at 1,700 g, 4˚C, for 5 min, the supernatant was collected and labelled as soluble fraction, and 100 µl sample loading buffer was added to the chromatin pellets for denaturing. Denatured proteins were loaded to the 4% to 12% gradient SDS-PAGE (GenScript) and western blot was performed as previous.

3.5.8. Real-Time Quantitative PCR (qPCR)

Purified total RNA (10 pg to 5 µg) was reverse transcribed into cDNA with the amfiRivert cDNA Synthesis Platinum Master Mix (GenDepot). Real-time quantitative PCR was performed as previous. All the primers were synthesized from Integrated DNA Technologies and listed in Table3.2.

Table 3.2.Primers for Real-Time Quantitative PCR (part2)

Target	Direction	Sequence (5'->3')
b-Actin	Forward	GACAGGATGCAGAAGGAGA
	Reverse	CCACATCTGCTGGAAGGTGG
TET1	Forward	GCAGCGTACAGGCCACCACT

Table 3.2. Primers for Real-Time Quantitative PCR (part2) (Continued)

Target	Direction	Sequence (5'->3')
TET1	Reverse	AGCCGGTCGGCCATTGGAAG
TET2	Forward	TTCGCAGAAGCAGCAGTGAAGAG
	Reverse	AGCCAGAGACAGCGGGATTCCTT
TET3	Forward	GACGAGAACATCGGCGGCGT
	Reverse	GTGGCAGCGGTTGGGCTTCT
DNMT1	Forward	CCCCTGAGCCCTACCGAAT
	Reverse	CTCGCTGGAGTGGACTTGTG
DNMT3A	Forward	TTCTACCGCCTCCTGCATGAT
	Reverse	GCGAGATGTCCCTCTTGTCACTA
DNMT3B	Forward	GAATTACTCACGCCCCAAGGA
	Reverse	ACCGTGAGATGTCCCTCTTGTC
Dnmt1	Forward	CCTAGTTCCGTGGCTAC AGGAGAA
	Reverse	TCTCTCTCCTCTGCAGCCGACTCA
Dnmt3a	Forward	GCCGAATTGTGTCTTGGTGGATGACA
	Reverse	CCTGGTGGAATGCACTGCAGAAGGA
Dnmt3b	Forward	GGATGTTCGAGAATGTTGTGG
	Reverse	GTGAGCAGCAGACACCTTGA
Tet1	Forward	GAGCCTGTTCCTCGATGTGG
	Reverse	CAAACCCACCTGAGGCTGTT

Table 3.2. Primers for Real-Time Quantitative PCR (part2) (Continued)

Target	Direction	Sequence (5'->3')
Tet2	Forward	TGCAAAACCTGGCTACTGTC
	Reverse	AACATGCAGTGACTCCTGAG
Tet3	Forward	TCCGGATTGAGAAGGTCATC
	Reverse	CCAGGCCAGGATCAAGATAA
Tbx20	Forward	TATTCAGCATACTCCTAC
	Reverse	GTTAGTCTTGTCAATACG
Nppa	Forward	CCTGCCTTCATCTATCAC
	Reverse	TGAGGATCTACCTTAATATGC
n-Myc	Forward	GACGAGGAGGAAGATGAA
	Reverse	CTTGTTGTTAGAGGAGGAA
hb-Actin	Forward	GAGAAGACTAGTGCCAACAGC
	Reverse	TGGGCATCAAAGTAGCCTTTA
Gapdh	Forward	GTGTTCCTACCCCCAATGTGT
	Reverse	ATTGTCATACCAGGAAATGAGCTT

3.5.9. RNA-Seq Library Construction and Data Analysis

Poly-A tailed messenger RNA was enriched with a Poly(A)Purist™ MAG Kit (Thermo Fisher Scientific). Enriched mRNA was used for RNA-seq library preparation by using a NEBNext® Ultra™ Directional RNA Library Prep Kit (NEB) according to the manufacturer's instructions. The quality of libraries was checked by an Agilent High

Sensitivity DNA kit (Agilent Technologies). The library was sequenced using an Illumina NextSeq 500 instrument (150 cycle, paired end).

RNA-seq data were mapped to mm10 genome assembly using tophat-2.1.1 with default parameters. Cufflinks and cuffdiff were used to call significantly differentially expressed genes (DEGs) (q-value <= 0.05) between WT and Tet2/3-DKO groups at E12.5 and E15.5 developmental stages. In-house R scripts were used to plot the scatter plot for DEGs. DEGs functional enrichment was performed using GSEA. RNA-seq data of mouse embryo hearts ranging from E10.5 to P0 were downloaded from ENCODE (https://www.encodeproject.org/). R package gplots was used to plot heatmaps for DEGs.

3.5.10. Single cell RNA-seq (scRNA-seq) library preparation and data analysis

Embryonic hearts were harvested and digested into single cells using 1 mg/ml collagenase I (Worthington). Single-cell RNA-seq libraries were generated using the Chromium Single-Cell 3' Reagent V2 Kit (10x Genomics) according to the manufacturer's protocol. Briefly, single cell GEM was generated and barcoded in a Chromium Controller (10x Genomics). Then RNA transcripts from single cells were reverse transcribed, amplified and fragmented. Library generation was finished by incorporating the adapter and sample indices into the fragmented cDNA. Agilent Bioanalyzer 2100 (Agilent) was used to profile the sizes of the pre-amplified cDNA and the libraries. Libraries were subjected to highthroughput sequencing on an Illumina NextSeq 500 system using the NextSeq 500 High Output v2 Kit (Illumina) with a customized paired end, dual indexing (26/8/0/58-bp) format as recommended by 10× Genomics. Single cell data analysis was performed as previous.

3.5.11. WGBS Library Construction and Data Analysis

Purified genomic DNA (with 5% of unmethylated lambda DNA spike-in, Promega) was sheared to till reaching a fragment size of 200~500 bp using Bioruptor UCD300 (Diagenode) according to manufacturer's instructions. Sheared DNA was ligated with methylated adaptors (NEBNext® Multiplex Oligos for Illumina®, NEB) by using a NEBNext® Ultra™ II DNA Library Prep Kit (NEB). Methylated adaptor-ligated DNA fragment was used for bisulfited conversion reaction with EZ DNA Methylation-Lightning Kit (Zymo Research), then bisulfite converted DNA was amplified using KAPA HiFi HotStart Uracil+ ReadyMix PCR Kit (Kapa Biosystems) with 8 cycles of PCR. Amplified DNA was purified by AMPureXP beads and examined by Agilent High Sensitivity DNA kit (Agilent Technologies) for quality check. Library concentration was determined by a Qubit 4 fluorometer (Thermo Fisher Scientific). Prepared libraries were sequenced using an Illumina NextSeq 500 instrument (150-cycle, paired-end).

Raw fastq files for WGBS (from E10.5-P0 stage) were downloaded from ENCODE. Raw fastq files were mapped to the hg19/mm10 genome assembly using bsmap-2.89 software with "-v 6 -n 1 -q 3 -r 1" parameters. The bisulfite conversion ratios were estimated using unmethylated lambda DNA. Mcall modual in MOABS was used to call the mCG/CG ratios for each CpG site. Mcomp modual was used to call DMRs with parameter "--minNominalDif=0.2 --minDmcsInDmr 3 --maxDistConsDmcs 500". The CpGs with coverage >= 5 was used for downstream analysis. The function prediction of DMRs was used for GREAT analysis. UCSC genome browser tracks were generated by using the Mmint ucsc.py function.

3.5.12. ATAC-seq Library Construction and Data Analysis

ATAC-seq library preparation was performed as described previously[133]. Briefly, 50,000 cells were collected in ice-cold PBS. Nuclei were isolated in Cold Lysis Buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL). The transposition reaction was performed by using a Nextera DNA Library Preparation Kit (Illumina) with modified tagmentation condition (37°C for 30 min). Tagmented DNA was purified by E.Z.N.A.® MicroElute Cycle Pure Kit (Omega BIO-TEK), then amplified with the KAPA real-time library amplification kit (Kapa Biosystems) followed by library purification using AMPure XP beads. The quality of purified DNA libraries was checked by Agilent High Sensitivity DNA kit (Agilent Technologies). The library was sequenced using an Illumina NextSeq 500 instrument (150 cycle, paired end) .

Bowtie2 with '-very-sensitive' option was used to map the high-quality reads to mm10 version of human genome. The uniquely properly paired mapped reads were extracted for downstream analysis. MACS2 with the '-nomodel' and '-extsize 147' was used to call ATAC peaks. Bedtools intersect (at least 1 bp overlap) was used to identify overlapped 5hmC peaks and ATAC peaks. We first cut the genome to 10kb equal size bins and use bigwigOverbed to calculate each bin's 5hmC and ATAC signals. R package geneplotter was used to plot the density scatterplots.

3.5.13. Cleavage Under Targets and Release Using Nuclease (Cut&Run)

Cut&Run was performed according to published protocol[119]. Briefly, single cells were attached to the concanavalin A-coated magnetic beads (Bangs Laboratories) followed by the *in-situ* binding of the antibody and pA-MN specifically to the target protein. Cleaved fragments were released after exposure to calcium. DNA was extracted from the

supernatant containing released chromatin fragments. Libraries were prepared using the ThruPLEX DNA-seq Kit from Rubicon Genomics (R400406) according to the manufacture's instruction with slightly modification of changing extension time of library amplification steps to 20 second. Libraries were sequenced using the NextSeq 500 High Output v2 Kit (Illumina, San Diego, CA) with a customized paired end, dual indexing (40/8/0/40-bp) format.

We analyzed CUT&RUN data using the script on github (https://github.com/Henikoff/Cut-and-Run). Briefly, we first mapped paired end raw fastq files to mm10 use bowtie 2.2.5 with parameters "--local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700". Then we used picard to remove duplication reads; bamToBed was used to transform bam file to bed file. Next, we used spike_in_calibration.csh script to perform spike in normalization. BedGraphtobigwig was used to transform the bedGraph file to bigwig file, which was used to perform the visualization.

3.5.14. HiChIP Library Construction and Data Analysis

HiChIP library preparation was performed as described previously[122]. Briefly, 1 million crosslinked cardiac cells from mouse embryos or 2 million mouse embryonic stem cells were lysed with Hi-C Lysis Buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% NP-40 with 1X protease inhibitor cocktail), followed by digestion with 150 U of MboI restriction enzyme (NEB) for 2 hours at 37°C. Fill-in master mix containing biotin-dATP (Thermo Fisher Scientific) was added to digest nuclei to generate enzyme-digested overhang and mark the DNA ends with biotin. Subsequently, T4 DNA ligase (NEB) was added to the reaction and incubated for 4 hours at room temperature to achieve proximity ligation. After

that, nuclei were resuspended with Nuclear Lysis Buffer (50 mM Tris-HCl pH7.5, 10 mM EDTA, 1% SDS with 1X protease inhibitor cocktail) and transferred to Covaris millitube for fragmentation. Fragmented samples were precleared by adding protein A/G beads (Thermo Fischer Scientific) for 1 hour at 4°C, and 2 µg Smc1a antibody was added with an overnight incubation. On day 2, protein A beads (Thermo Fisher Scientific) were added to the reaction for 2 hours to capture the beads. After bead capturing, the samples were washed three times each with low salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 7.5, 150 mM NaCl), high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 7.5, 500 mM NaCl), and LiCl wash buffer (10 mM Tris-HCl pH 7.5, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1mM EDTA) at room temperature. After these steps, ChIP samples were resuspended in a DNA elution buffer (50 mM sodium bicarbonate pH8.0, 1% SDS) and incubated for 10 minutes at room temperature, followed by shaking for 3 minutes at 37°C. DNA eluted from the beads were collected twice, followed by reverse crosslinking. Reverse crosslinked DNA were then purified by using an E.Z.N.A.® MicroElute Cycle Pure Kit (Omega BIO-TEK). Purified samples were used for biotin pull-down. Resuspended Streptavidin C-1 (Thermo Fisher Scientific) with 2X Biotin binding buffer (10 mM Tris-HCl pH7.5, 1 mM EDTA, 2 M NaCl) was added to the samples and proceed biotin capturing procedure by incubating 15 min at room temperature, followed by washes with Tween wash buffer (5 mM Tris-HCl pH7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20) and 1X TD buffer (10 mM Tris-HCl pH7.5, 5 mM magnesium chloride, 10% dimethylformamide), respectively. After wash, on-bead tagmentation by using Tn5

transposase (Illumina) was performed for 10 minutes at 55°C with interval shaking, followed by several washes with 50 mM EDTA, Tween-20 wash buffer, and 10 mM Tris-HCl respectively. After wash, the reaction beads were resuspended in a PCR master mix (Q5® High-Fidelity 2X Master Mix, NEB, with Nextera Ad1.1 (Universal) and Ad2.X (barcoded) primers) for library amplification. Amplified on-bead DNA were eluted using a magnet and purified with an E.Z.N.A.® MicroElute Cycle Pure Kit (Omega BIO-TEK). The quality of libraries was checked by Agilent High Sensitivity DNA kit (Agilent Technologies). The library was sequenced on an Illumina NextSeq 500 instrument (150 cycle, paired end).

1 million cells from each condition were used to perform HiChIP experiments. To improve the statistic power, we merged the two biological replicates to increase sequencing depth. HiC-Pro was used to map the raw paired-end fastq files to mm10 genome assembly and identify the uniquely validated paired reads. build_raw_maps.sh and ice_norm.sh embeded in HiC-Pro pipeline were used to generate the raw contact map and normalized contact map. Fit-Hi-C was used to identify the significant Paired End Tags (PETs) between any two bins (5 kb) with a p value less than 0.05. We linked the PETs with the whole genic regions. The PETs linked to specific genes were counted and normalized as shown in a previous study. The RNA-seq, CMS-IP, ATAC-seq, ChIP-seq data and pairwise files containing PETs information were uploaded to WashU Epigenome Browser (http://epigenomegateway.wustl.edu)

3.5.15. Prediction of AB Compartment

The output file *_allValidPairs.hic from HiC-Pro pipeline was used as input file for juicer_tools.1.7.5_linux_x64_jcuda.0.8.jar eigenvector function[134]. The

*_allValidPairs.hic file stores all the raw interaction paired reads between any two genomic bins from the same chromosome. The eigenvector for each chromosome with KR normalization at 50 kb resolution were calculated. For each chromosome, we manually checked the overlap between the compartment assignment and the accessible regions from our ATAC-seq data to decide if the compartment assignment need to be flipped.

4. TET MEDIATED EPIGENETIC REGULATION IS IMPORTANT FOR SECOND HEART FIELD DEVELOPMENT

4.1. Summary

Our previous studies suggested that Tet2 and Tet3 deletion resulted in ventricular non-compaction cardiomyopathy (NCC) from embryonic day 13.5, which indicates Tet protein plays important role in cardiac development. However, it is unclear which cell population is directly affected by Tet loss-of-function during cardiac development. In order to address this question, we applied advanced single-cell RNA-seq technology in heart tissue purified from Tet1-3 triple floxed mice with Nkx2.5Cre or Isl1MerCreMer. Our analysis suggested that Tet mediated epigenetic regulation is important for cardiac progenitor differentiation toward cardiomyocytes at second heart field. Deletion of Tet protein impaired the epithelial to mesenchymal transition (EMT) of the second heart field (SHF) progenitors during heart development, which caused the accumulation and myocyte lineage specification defect of the SHF progenitors.

4.2. Introduction

Our previous study showed that specific knock-out of Tet2 and Tet3 during heart development could disrupt the YY1 chromatin binding and YY1 mediated long range enhancer and promoter interactions, which will lead to non-compaction cardiomyopathy during mouse embryonic heart development[39]. Depletion of Tet2 and Tet3 also affected zebrafish heart development by blocking the active demethylation of *Inhbaa* gene in endocardium and Sox9b gene in myocardium, which resulted in impairing the migration of proepicardial cells[81]. However, it is still unclear which cell population is directly

affected by Tet loss-of-function during cardiac development. In this study, we applied single-cell RNA-seq (scRNA-seq) technology in heart tissue purified from Tet1-3 triple floxed mice crossed with cardiac development specific Nkx2.5Cre[128] and cardiac second heart field specific Isl1MerCreMer[135] mice. Both Nkx2.5Cre and Isl1MerCreMer mediated Tet-TKO embryonic hearts showed shorter outflow tracts at E9.5 and developed other second heart related structure defects, such as ASD and VSD, at later E15.5. Our single cell transcriptome analysis showed the increase of the second heart field progenitors and decreased of the cardiomyocytes in Tet-TKO embryonic hearts. Epithelial to mesenchymal transition was the most affected pathway after Tet complete depletion. We concluded that Tet-TKO could disrupt the epithelial to mesenchymal transition of second heart field progenitors during heart development, which caused the SHF development defects.

4.3. Results

4.3.1. Complete Depletion of Tet Enzymes Caused the Second Heart Field Development Defects

In order to study how complete loss of Tet enzymes affects mouse heart development, we generated the cardiac development specific Tet triple knock-out (Tet-TKO) mice models (Figure 4.1A). Cardiac progenitor specific expression Cre recombinases (Nkx2.5Cre[128] and Isl1MerCreMer[135]) mice lines were crossed with Tet triple flox mice. Nkx2.5Cre mediated the Tet-TKO in the cardiac progenitors from both the FHF and the SHF, while Isl1MerCreMer mediated the Tet-TKO only in the SHF progenitors after tamoxifen treatment. A loxP-flanked STOP sequence followed by the enhanced yellow fluorescent protein gene (EYFP)[136] was also crossed into the genome of Tet-TKO mice. Tet-TKO cell would be labelled by EYFP signal because of the presence of Cre recombinases (Figure

4.1A). We clearly observed the EYFP signal overall the embryonic hearts in E9.5 Nkx2.5Cre embryos and only observed the EYFP signals in the SHF related structures in the E9.5 Isl1MerCreMer embryonic hearts, which also indicated the correct function of Cre recombinases (data not shown). In order to further confirm the depletion of Tet enzymes, 5hmC immunostaining was performed on the E9.5 Control and Tet-TKO hearts from both Nkx2.5Cre and Isl1MerCreMer embryos. 5hmC is the major product of Tet enzymes. The immunostaining results showed that 5hmC in the hearts of Tet-TKO embryos was significantly decreased when compared to the control (Figure4.1 B, C).

Both Nkx2.5Cre and Isl1MerCreMer mediated Tet-TKO mice were embryonic lethal. We performed timed mating to get the Tet-TKO embryos. Images of embryo morphology showed that both Nkx2.5Cre and Isl1MerCre mediated Tet-TKO hearts had significant shorter outflow tracts at E9.5(data not shown). Comprehensive histological analysis was conducted to evaluate the heart development defects (Figure4.1D-G). In both Nkx2.5Cre and Isl1MerCreMer mediated Tet-TKO, significant shorter OFTs were also observed from the H&E results of E9.5 embryos, which were consistent with the observations of E9.5 embryo images (Figure4.1D-G). This indicated that Tet-TKO might cause the second heart development defects. H&E analysis of control and Tet-TKO embryos further showed the second heart field related structure defects, such as atrial septal defect (ASD) and ventricular septal defect (VSD), in Tet-TKO embryos at later E15.5 stage (Figure4.1D, F). All these results suggested that Tet complete depletion caused the second heart field development defects during heart development[137-139].

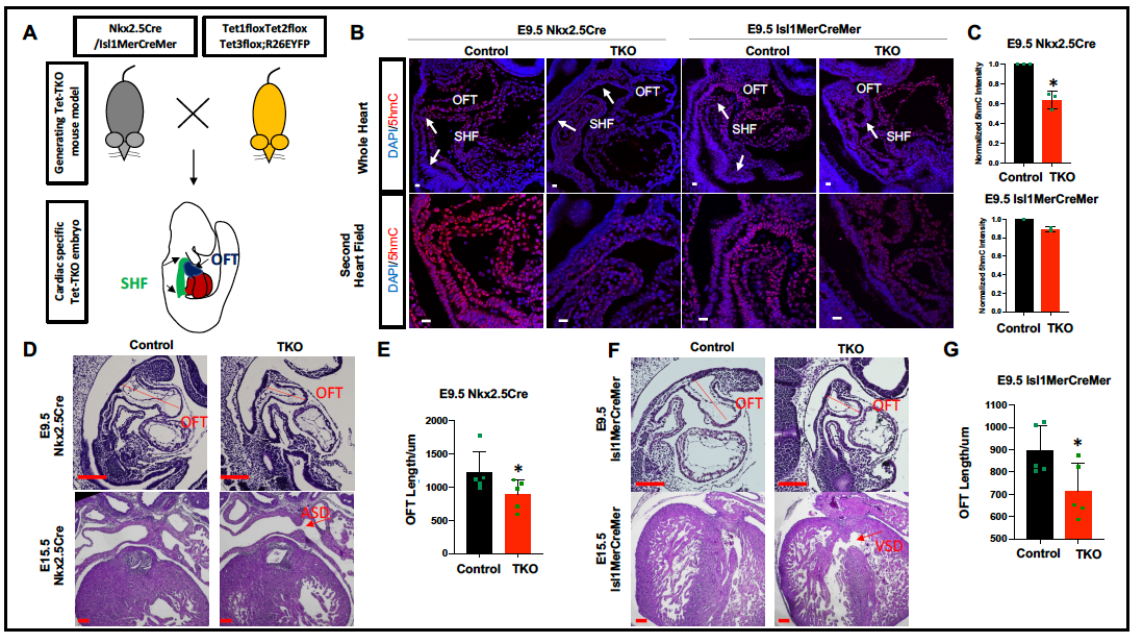

Figure 4.1.Complete Depletion of Tet Enzymes Caused the Second Heart Field Development Defects.

A) Illustration of the strategy to generate the cardiac development specific Tet triple knock-out mice. Cardiac progenitor specific Nkx2.5Cre and second heart field specific Isl1MerCreMer were used; B) Representative 5hmC immunostaining images of control and Tet-TKO embryonic hearts in E9.5 Nkx2.5Cre and Isl1MerCreMer embryos. Scale bar, 20um; C) (Upper)Quantification of relative 5hmC intensity from E9.5 Nkx2.5Cre control and Tet-TKO embryonic hearts 5hmC immunostaining images. (Lower) Quantification of relative 5hmC intensity from E9.5 Isl1MerCreMer control and Tet-TKO embryonic hearts 5hmC immunostaining images. Quantification was performed by Image J from three independent experiments. Data was shown as mean ± S.D. * $p<0.05$ (compared to control, two-tailed student's t-test); D) (Left)Representative H&E staining images of Nkx2.5Cre control and Tet-TKO embryonic hearts from E9.5 and E15.5. Scale bar, 500um. (Right) Quantification of the OFT length from the E9.5 Nkx2.5Cre control

and Tet-TKO H&E staining images. Quantification was performed by ImageJ from five independent experiments. Data was shown as mean ± S.D. * $p<0.05$ (compared to control, two-tailed student's t-test). ASD, atrial septal defect; E) (Left)Representative H&E staining images of Isl1MerCreMer control and Tet-TKO embryonic hearts from E9.5 and E15.5. Scale bar, 500um. (Right) Quantification of the OFT length from the E9.5 Isl1MerCreMer Control and Tet-TKO H&E images. Quantification was performed by ImageJ from five independent experiments. Data was shown as mean ± S.D. * $p<0.05$ (compared to control, two-tailed student's t-test). VSD, ventricular septal defect.

4.3.2. Tet Triple Knock-Out Caused the Defects of Second Heart Field Progenitors During Heart Development

Since the OFTs in Tet-TKO embryos were shorter when compared with control embryonic hearts, we asked if the proliferation or cell death were abnormal in the Tet-TKO embryonic hearts. We performed immunostaining of the cell proliferating markers (Ki67 and PHH3) and cell apoptosis marker (Cleaved caspase3) but we didn't find any significant differences between control and Tet-TKO hearts (data not shown). In order to uncover the cellular defects and the mechanisms, we performed the single cell RNA-seq using E9.5 control and Tet-TKO hearts from Nkx2.5Cre embryos (Figure 4.2A). We collected a total of 8,085 cells from both control and Tet-TKO hearts (Figure 4.2B). After applying the UMAP to cluster the cells from control and Tet-TKO hearts based on their transcriptomes, we found the transcriptomes between control and Tet-TKO hearts were dramatic different (Figure 4.2B). Based on the distinct gene expressions of each cluster, we identified 10 different clusters and assigned specific cell type to each cluster (Figure 4.2C). The percentages of each cell type in control and Tet-TKO were calculated. Interestingly, we

found that cardiomyocytes population, as well as the endothelia cells, was dramatically decreased in Tet-TKO heart; and cardiac progenitors, neural crest and epithelial populations were dramatically increased in Tet-TKO heart (Figure 4.2D). These results indicated that Tet triple knock-out might disrupt the cell composition in mouse embryonic heart during development.

Cell composition was complicated in embryonic heart. It contained almost all cell types during heart development. we performed the pseudo-time analysis[90] **to compare the dynamic differentiation ability between control and Tet-TKO hearts**. We observed distinct trajectories between control and Tet-TKO groups. In control heart progenitors were clustered together and differentiated into myocyte lineage. However, progenitors were more heterogeneous and less differentiating into myocyte lineages in Tet-TKO heart (Figure 4.2E). This indicated that Tet-TKO would block the myocyte lineage differentiation of the cardiac progenitors. Our progenitor and cardiomyocyte immunostaining results confirmed our pseudo-time analysis in single cell RNA-seq dataset (Figure 4.2F, G). Isl1 is the marker of second heart field progenitors[140] and immunostaining result showed the increase of Isl1 positive progenitors in the OFT of the Tet-TKO embryonic heart (Figure 4.2 F). a-Actinin is the marker of cardiomyocytes[141] and our results showed the decrease of a-Actinin staining positive cells in Tet-TKO embryonic hearts. Taken together, we concluded that Tet triple knock-out caused the second heart field progenitors lineage specification defects during heart development.

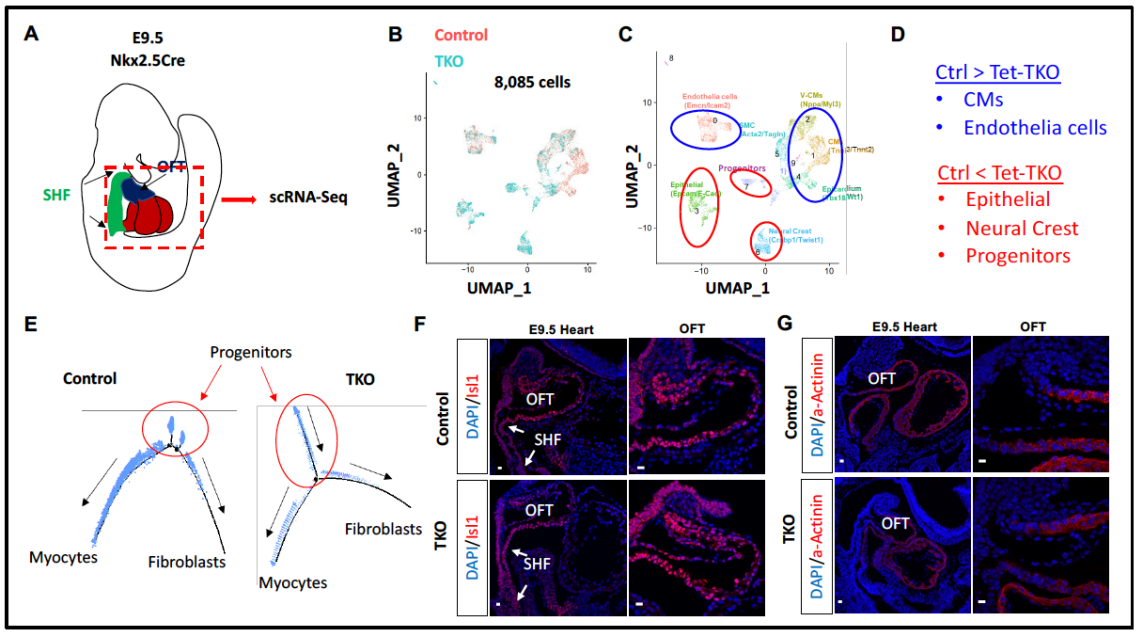

Figure 4.2.Tet Triple Knock-Out Blocks Second Heart Field Progenitors Differentiating to Myocytes.

A) Illustration of the strategy to collect samples from E9.5 Nkx2.5Cre control and Tet-TKO hearts for single cell RNA analysis; B) the UMAP clustering plot of E9.5 Nkx2.5Cre control and Tet-TKO scRNA-seq; Red color labelled cells are from control heart and green color labelled cells are from Tet-TKO heart; C) the UMAP plot of cell types from E9.5 Nkx2.5Cre control and Tet-TKO scRNA-seq; There are a total of 10 cell types. SMC, smooth muscle cells; V-CM, ventricular cardiomyocytes; CM, cardiomyocytes; D) Comparison of the cell type percentages. (Upper) Significant decreased cell types in Tet-TKO heart (blue color); (Lower) Significant increased cell types in Tet-TKO heart (red color); E) Pseudo-time analysis of E9.5 Nkx2.5Cre control and Tet-TKO scRNA-seq. (Left) the trajectory of control heart; (Right) the trajectory of Tet-TKO heart; F) Representative Isl1 immunostaining images of control and Tet-TKO embryonic hearts in E9.5 Nkx2.5Cre embryos. Scale bar, 20 μm. OFT, outflow tract; SHF, second heart field;

G) Representative a-Actinin immunostaining images of control and Tet-TKO embryonic hearts in E9.5 Nkx2.5Cre embryos. Scale bar, 20 μm. OFT, outflow tract; SHF, second heart field.

4.3.3. Tet Triple Knock-Out Affected the Epithelial to Mesenchymal Transition (EMT) of SHF Progenitors During Heart Development

In order to fully take advantages of scRNA-seq dataset, we performed the RNA velocity analysis for the E9.5 Control and Tet-TKO datasets (Figure 4.3A). We found that Isl1 positive second heart field progenitors had the trend of differentiating into other downstream cell type lineages, including myocyte lineage, in the control heart. However, second heart field progenitors in the Tet-TKO heart had limited ability to differentiate and cardiomyocyte lineage differentiation was impaired (Figure 4.3A).

Although scRNA-seq was a powerful tool to study the transcriptome at single cell level, its disadvantage was the lower sequencing depth, which made it difficult to get enough reads for solid downstream analysis[142]. In order to overcome this disadvantage, we harvested YFP positive cells from the E9.5 Nkx2.5Cre control and Tet-TKO hearts by flow cytometry sorting and performed the bulk RNA-seq experiments for the YFP positive cell populations. The YFP positive cell percentages in E9.5 control and Tet-TKO hearts were nearly 50% (data not shown). Data analysis of our YFP positive RNA-seq datasets provided the significantly enriched gene sets in both control and Tet-TKO groups (Figure 4.3 B, C). The top3 enriched gene sets in control heart were oxidative phosphorylation, myogenesis and fatty acid metabolism, which were all myocytes related (Figure 4.3B). This was consistent with functions of Nkx2.5, a cardiac specific transcriptional factor[143-145]. Progenitors related gene sets were highly enriched In the Tet-TKO heart. The top3

enriched gene sets were hedgehog signaling, epithelial-mesenchymal transition and mitotic spindle (Figure 4.3C). Epithelial-mesenchymal transition played important roles during second heart field progenitor differentiating to cardiomyocytes[146-148]. Our previous scRNA-seq results indicated the second heart field progenitor differentiation defect in the Tet-TKO heart (Figure 4.2D, E). scRNA-seq result also showed the increase of epithelial cells in Tet-TKO hearts (Figure 4.3D). Taken together, we concluded that Tet complete depletion might impair the epithelial mesenchymal transition of SHF progenitors during heart development.

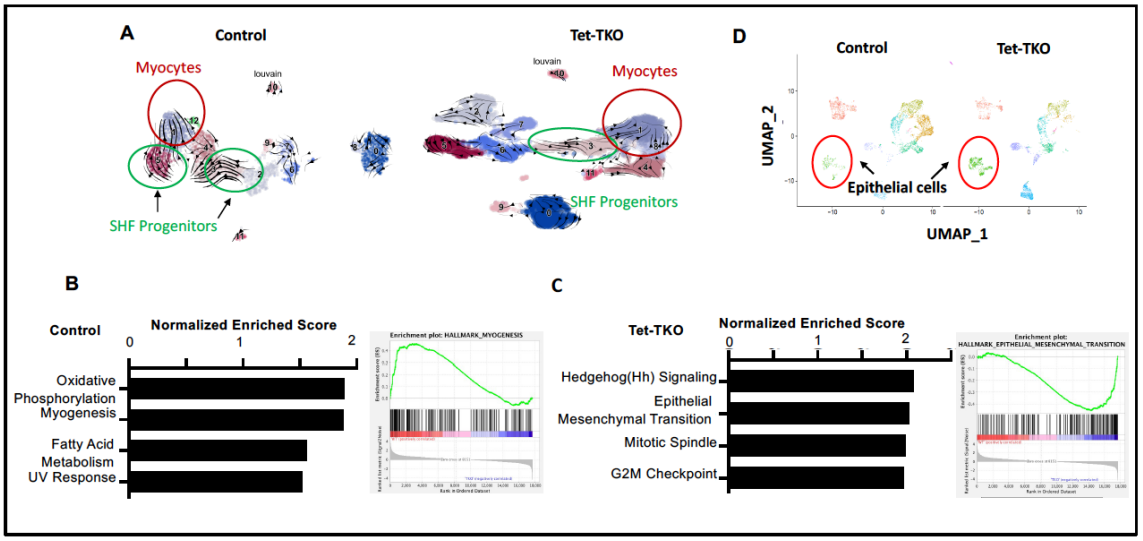

Figure 4.3.Tet-TKO Affected the Epithelial-Mesenchymal Transition of SHF Progenitors.

A) RNA velocity analysis of E9.5 Nkx2.5Cre Control and Tet-TKO scRNA-seq; (Left) control group; (Right) Tet-TKO group; Red circle labels the myocytes and green circle labels the SHF progenitors; SHF, second heart field; B) Gene set enrichment analysis of RNA-seq dataset from YFP positive cells in E9.5 Nkx2.5Cre control embryonic heart. (Left) Significant enriched gene sets; (Right) Myogenesis gene set; C) Gene set

enrichment analysis of RNA-seq dataset from YFP positive cells in E9.5 Nkx2.5Cre Tet-TKO embryonic heart. (Left) Significant enriched gene sets; (Right) Epithelial-mesenchymal transition gene set; D) UMAP plot of epithelial population in E9.5 Nkx2.5Cre Control and Tet-TKO scRNA-seq;(Left) Control group; (Right)Tet-TKO group.

4.3.4. EMT Related Transcriptional Factors Might Be the Target of Tet Enzymes

In order to test if the EMT was disrupted after Tet depletion during heart development, we performed immunostaining for both epithelial marker(E-Cadherin)[149] and mesenchymal marker(N-Cadherin)[150] in E9.5 Nkx2.5Cre control and Tet-TKO hearts(Figure 4.4 A, B). Our results indicated that epithelial cell population was increased in the second heart field area of the Tet-TKO embryonic heart (Figure 4.4A). Mesenchymal cell population was decreased in the second heart field area in the Tet -TKO heart (Figure 4.4B). This suggested that epithelial-mesenchymal transition of second heart field progenitors was indeed affected after Tet complete depletion during heart development. In order to further uncover the underlying mechanism, we used the Tet triple knock-out mouse embryonic stem cells cardiac differentiation model that we established before. We differentiated two Tet-TKO clones with WT mESCs and analyzed the EMT at day 5 of differentiation. Our real-time quantitative PCR results showed that epithelial cell marker, E-cadherin, was increased in differentiated Tet-TKO group, while both mesenchymal markers (N-Cadherin and Vim) and EMT related transcriptional factors (Snai2 and Twist1) [150,151]were downregulated in Tet-TKO group (Figure 4.4 D). This suggested that Tet triple knock-out could downregulate the EMT regulators, such as Snai2 and Twist1, to impair the epithelial-mesenchymal transition during heart development.

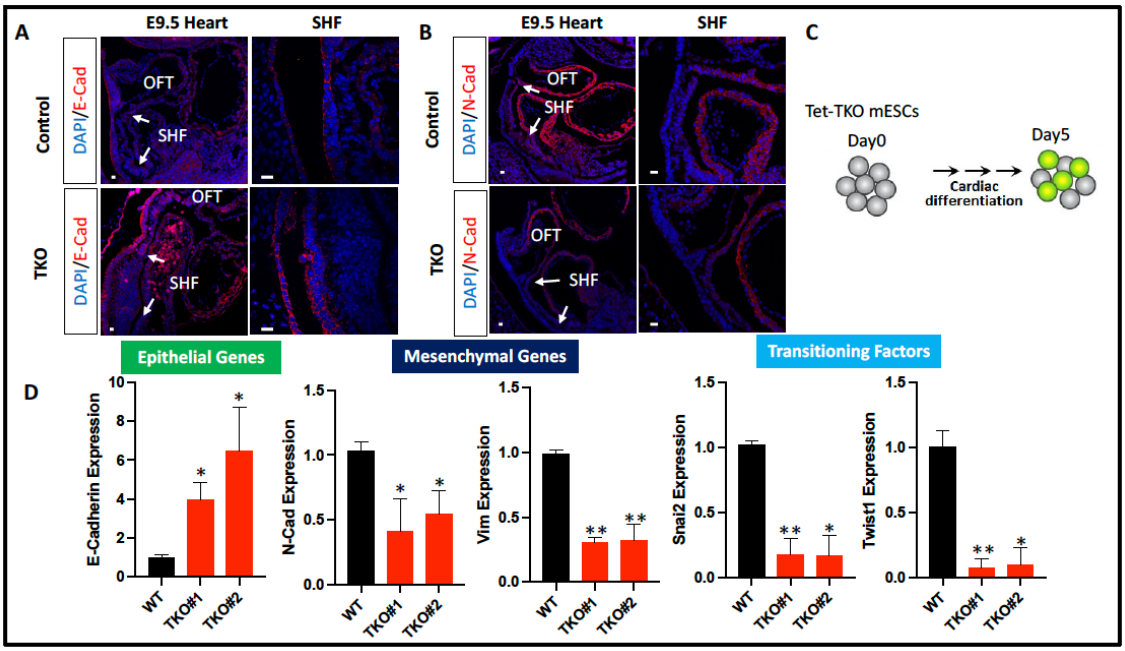

Figure 4.4.EMT Related Transcriptional Factors Might Be the Target of Tet Enzymes.

A) Representative E-Cadherin immunostaining images of control and Tet-TKO embryonic hearts in E9.5 Nkx2.5Cre. Scale bar, 20 µm. B) Representative N-Cadherin immunostaining images of Control and Tet-TKO embryonic hearts in E9.5 Nkx2.5Cre. Scale bar, 20 µm. C) Illustrator of the mESCs cardiac differentiation model; D) Real-time qPCR analysis of expression of epithelial-mesenchymal transition factors in the control and TKO groups (two clones) at day 5 of differentiation. Data were shown as mean ± S.D; n= 3. ** $p < 0.01$, * $p < 0.05$ (compared to the WT; two-tailed Student's t-test).

4.4. Discussion

Our previous studies showed cardiac specific knock-out of Tet2 and Tet3 could affect the chromatin binding of critical transcriptional factors, disrupted higher-order chromatin structures and lead to non-compaction cardiomyopathy during mouse heart development[39]. This study indicated that cardiac development specifically complete

depletion of three Tet enzymes would cause the second heart field progenitors differentiation defects. Tet triple knock-out could affect the epithelial-mesenchymal transition of the second heart field progenitors and impaired the cardiac differentiation, which was the first study to show how Tet enzymes regulated the cardiac lineage specification during early heart development.

Our study illustrated the function of Tet enzymes in second heart field development. Complete knock-out of Tet enzymes could cause the second heart development defects, which are also known as the conotruncal defects (CTDs). CTDs are the major disease types of congenital heart diseases (CHDs), the most common birth defect in human. They accounts nearly 30% of CHD patients[152]. Previous studies reported hypoxia was one of the factors that leads to CTDs[153,154]. Considering oxygen is the important substrate of Tet enzymes mediated active DNA demethylation[22], our study may provide the molecular evidence to understand how hypoxia could cause the conotruncal defect in human, which may help to prevent and treat the congenital heart diseases.

Taking advantage of the mESCs cardiac differentiation model, we revealed that Tet enzymes downregulated transcriptional factors, such as Snai2 and Twist1, which played critical roles in inducing epithelial-mesenchymal transition. Those factors could be the direct targets of Tet proteins, but more evidence should be present to confirm this. This will be one of our following research directions.

4.5. Materials and Methods

4.5.1. Animal Models

Animal studies were approved by the Institutional Animal Care Use Committee (IACUC) of the Institute of Biosciences and Technology, Texas A&M University. Most mouse

strains bear a C57BL/6 genetic background unless otherwise noted. Tet1f/f[155], Tet2f/f[127], Tet3f/f3[70], Nkx2.5-Cre (The Jackson Laboratory 024637)[128], R26R-EYFP(The Jackson Laboratory006148)[136], and Isl1MerCreMer(The Jackson Laboratory 029566)[135] mouse strains were reported previously. Our mice breeding strategy was to generate the Tet triple flox mice with R26R-EFP first. Then Tet triple flox mice were mating with Cre mice to produce the cardiac development specific Tet triple deletion mice.

Timed mating and genotyping experiments were conducted by following the protocols in 3.5.2. Tet2, Tet3, Nkx2.5Cre genotyping primers were in Table3.1. Tet1, EYFP and Isl1MerCreMer genotyping primers were in the Table 4.1.

Table 4.1.Genotyping primers of Tet1, EYFP and Isl1MerCreMer Mice

Target	Direction	Sequences (5'->3')
Tet1WT	Forward	CAGTTCTATTCAGTAAGTAAGTGTGCC
	Reverse	GGTTGTGTTAAAGTGAGTTGCAAGCG
EYFP	Forward	AGGGCGAGGAGCTGTTCA
	Reverse	TGAAGTCGATGCCCTTCAG
Isl1MerCreMer	Wild type reverse	TAGAGTTTGGGAGGGCACAG
	Common	GCCACTATTTGCCACCTAGC
	Mutant reverse	CGGTTCAGCATCCAACAAG

4.5.2. Antibodies

For immunostaining: 5hmC (Active Motiff 39769, 1:1000), Isl1(Developmental Studies Hybridoma Bank clone 39.4D5, 1:100), a-Actinin (Abcam ab68167, 1:400), E-Cadherin (Cell Signaling Technology #5296, 1:200), N-Cadherin (GeneTex 127345, 1:100), Alexa

Fluor 568 goat anti-rabbit (Thermo Fisher Scientific A-11011, 1:1,000), Alexa Fluor 568 goat anti-mouse (Thermo Fisher Scientific A-11011, 1:1,000).

4.5.3. Paraffin Section of Embryos

All mouse embryos were dissected in phosphates buffered saline (PBS) using a Nikon SMZ800N dissecting microscopy. Embryos were then fixed overnight in 4% PFA, dehydrated with graded ethanol and embedded in paraffin. Sections were cut at the thickness of 7 μm using the microtome (Olympus CUT 4060). Slides were dried at 37 ˚C overnight for following hematoxylin-eosin (H&E) staining and Immunofluorescence (IF) staining.

4.5.4. Hematoxylin-Eosin(H&E) Staining and Quantification

H&E staining was performed as previously described[130,131]. In briefly, paraffin sections were incubated in xylene to dissolve all the wax and then rehydration in a graded series of ethanol (100% to 70%). After hematoxylin-eosin staining, another graded series of ethanol (70% to 100%) were applied again to dehydrate sections and xylene was used to dissolve the ethanol. After mounting with Poly-Mount (Polysciences) medium, images of the sections were taken using a Nikon Eclipse Ci microscopy.

H&E images were quantified using ImageJ software. Sections from similar position were chosen for quantification. In order to quantify the length of the outflow tract, we drew a line in the middle of the OFT and quantify the lengthen. A total of five embryos were quantified from each group.

4.5.5. Immunofluorescence (IF) Staining

IF staining was performed following previously described protocols[131]. Briefly, tissue sections were dewaxed in xylene rehydrated in a graded series of ethanol (100% to 70%). The antigens were retrieved by boiling sections in antigen retrieval buffer (Vector

Laboratories) for 20 min. Sections were permeabilized in 0.2% Triton X-100 in PBS for 15 min and incubated in 5% normal goat serum (Thermo Fisher Scientific) to block the unspecific antigens. After primary and 2nd antibodies incubation, DAPI (Biolegend #422801) was used to co-stain the nuclei. IF stained sections were imaged using a Nikon A1 confocal microscope or the Nikon W1 spinning microscope.

For 5hmC staining, sections were treated with 2N HCl for 30 min to expose the epitopes and then neutralized in 100 mM Tris-HCl (pH 8.5) for 10min before blocking with 5% normal goat serum.

For Isl1 staining, tyramide signal amplification (TSA) kit (AKOYA Biosciences NEL741001KT) was used to amplify the signal. Experiments were conducted following the manual. After the primary antibody incubation, 3% H_2O_2 was used to quench the endogenous peroxidase. Sections were then incubated with the HRP conjugated anti-mouse secondary antibody. Lastly, sections were incubated with the TSA working solution to add the appropriate chromogen to the Isl1, primary antibody and secondary antibody complex.

4.5.6. Real-Time Quantitative PCR (qPCR)

qPCR experiments were performed as in 3.5.8. Primers used were listed in Table4.2.

Table 4.2.Primers for Real-Time Quantitative PCR (Part3)

Target	Direction	Sequences (5'->3')
E-Cadherin	Forward	ACCATCACAGTGAAGCGG
	Reverse	TGGGTACACGCTGGGAAA
N-Cadherin	Forward	GGCAATCCCACTTATGGC
	Reverse	TCCGTGACAGTTAGGTTG

Table 4.2. Primers for Real-Time Quantitative PCR (Part3) (Continued)

Target	Direction	Sequences (5'->3')
Vim	Forward	AAGAGATGGCTCGTCACCTT
	Reverse	GGGTGTCAACCAGAGGAAGT
Snai2	Forward	TGGTCAAGAAACATTTCAACGCC
	Reverse	GGTGAGGATCTCTGGTTTTGGTA
Twist1	Forward	CGCTACAGCAAGAAATCGAGC
	Reverse	GCTGAGCTTGTCAGAGGG

4.5.7. Single Cell RNA-Seq Libraries Preparation and Data Analysis

E9.5 embryonic hearts from control and Tet-TKO groups were harvested for the single cell RNA-seq experiments. Libraries preparation, sequencing and data analysis were following previous description (2.5.11 and 3.5.10).

4.5.8. RNA-seq Libraries Preparation and Data Analysis

E9.5 embryonic hearts from control and Tet-TKO groups were harvested under the dissecting microscopy (Nikon SMZ800N) and single cell solution preparation was following the previous description (3.5.10). YFP positive cells were collected by flow cytometry. RNA was isolated using the AllPrep DNA/RNA Mini Kit (Qiagen) and the RNA-seq library preparation was using the SMART-Seq v4 kit (TaKaRa). RNA-seq libraries sequencing setting and data analysis were following the protocol described previously (3.5.9).

5. CONCLUSIONS

The ultimate goal of this study is to reveal the functions of Tet enzymes in mammalian heart development. To achieve this, we firstly took advantages of the CRISPR/Cas9 technology to generate Tet deficient mESCs and studied the functions of Tet enzymes during mESCs cardiac differentiation. In parallelly, we generated the cardiac development specific Tet deficient mouse models using the Cre-loxP system and investigated the functions of Tet enzymes during mouse heart development. We applied various epigenome approaches to further investigate the underlying molecular mechanisms.

Our study revealed that Tet mediated epigenetic regulation plays critical roles in orchestrating mESCs cardiac differentiation. It controls the methylation status of the regulatory elements, which are crucial for regulating the expression of cardiac development transcriptional factors, such as Hand1. Our study also revealed that Tet proteins play indispensable roles for mouse embryonic heart development. Cardiac specific Tet deficiency during embryonic development causes various congenital heart defects, such as non-compaction cardiomyopathy, short outflow tract, and atrial and ventricular septal defects. We uncovered that Tet mediated DNA demethylation associates with chromatin accessibility and regulates higher-order chromatin structures, which plays crucial roles in orchestrating cardiac specific transcriptional networks during embryonic heart development.

Altogether, our study illustrates that Tet proteins are crucial for embryonic heart development and Tet deficiency leads to congenital heart defects. Congenital heart defects

are the most common birth defects in human and affect 1 in 100 births. Our study provides critical epigenetic evidence for the prevention and treatment of congenital heart diseases.

6. REFERENCES

1 Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* **18**, 517-534, doi:10.1038/nrg.2017.33 (2017).

2 Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-220, doi:10.1038/nrg3354 (2013).

3 Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33 Suppl**, 245-254, doi:10.1038/ng1089 (2003).

4 Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* **6**, a019133, doi:10.1101/cshperspect.a019133 (2014).

5 Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089-1093, doi:10.1126/science.1063443 (2001).

6 Lyko, F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet* **19**, 81-92, doi:10.1038/nrg.2017.80 (2018).

7 Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760-1764, doi:10.1126/science.1147939 (2007).

8 Hermann, A., Goyal, R. & Jeltsch, A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* **279**, 48350-48359, doi:10.1074/jbc.M403427200 (2004).

9 Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-232 (1975).

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

11 Brabson, J. P., Leesang, T., Mohammad, S. & Cimmino, L. Epigenetic Regulation of Genomic Stability by Vitamin C. *Front Genet* **12**, 675780, doi:10.3389/fgene.2021.675780 (2021).

12 Ghazimoradi, M. H. & Farivar, S. The role of DNA demethylation in induction of stem cells. *Prog Biophys Mol Biol* **153**, 17-22, doi:10.1016/j.pbiomolbio.2019.12.005 (2020).

13 Lio, C. J. & Rao, A. TET Enzymes and 5hmC in Adaptive and Innate Immune Systems. *Front Immunol* **10**, 210, doi:10.3389/fimmu.2019.00210 (2019).

14 Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929-930, doi:1169786 [pii]10.1126/science.1169786 (2009).

15 Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935, doi:1170116 [pii]10.1126/science.1170116 (2009).

16 Weber, A. R. *et al.* Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat Commun* **7**, 10806, doi:10.1038/ncomms10806 (2016).

17 Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133, doi:10.1038/nature09303 [pii] (2010).

18 Huang, Y. & Rao, A. Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet* **30**, 464-474, doi:10.1016/j.tig.2014.07.005 (2014).

19 Hu, L. *et al.* Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* **155**, 1545-1555, doi:10.1016/j.cell.2013.11.020S0092-8674(13)01469-4 [pii] (2013).

20 Pastor, W. A., Aravind, L. & Rao, A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14**, 341-356, doi:10.1038/nrm3589 (2013).

21 Ko, M. *et al.* Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**, 122-126, doi:10.1038/nature12052 [pii] (2013).

22 Lian, C. G. *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* **150**, 1135-1146, doi:10.1016/j.cell.2012.07.033S0092-8674(12)01012-4 [pii] (2012).

23 Losman, J. A. & Kaelin, W. G., Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes Dev* **27**, 836-852, doi:10.1101/gad.217406.113 (2013).

24 Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17-30, doi:10.1016/j.ccr.2010.12.014 (2011).

25 Zhao, B. *et al.* Redox-active quinones induces genome-wide DNA methylation changes by an iron-mediated and Tet-dependent mechanism. *Nucleic Acids Res* **42**, 1593-1605, doi:10.1093/nar/gkt1090 (2014).

26 Blaschke, K. *et al.* Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222-226, doi:10.1038/nature12362[pii] (2013).

27 Minor, E. A., Court, B. L., Young, J. I. & Wang, G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem* **288**, 13669-13674, doi:10.1074/jbc.C113.464800 [pii] (2013).

28 Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem Sci* **36**, 7-18, doi:10.1016/j.tibs.2010.07.002S0968-0004(10)00134-9 [pii] (2011).

29 Rose, N. R., McDonough, M. A., King, O. N., Kawamura, A. & Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev* **40**, 4364-4397, doi:10.1039/c0cs00203h (2011).

30 Fu, X. *et al.* MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. *Proc Natl Acad Sci U S A* **110**, 17892-17897, doi:10.1073/pnas.1317397110 (2013).

31 Lv, X., Jiang, H., Liu, Y., Lei, X. & Jiao, J. MicroRNA-15b promotes neurogenesis and inhibits neural progenitor proliferation by directly repressing TET3 during early neocortical development. *EMBO Rep* **15**, 1305-1314, doi:10.15252/embr.201438923 (2014).

32 Song, S. J. *et al.* The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. *Cell Stem Cell* **13**, 87-101, doi:10.1016/j.stem.2013.06.003 (2013).

33 Song, S. J. *et al.* MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* **154**, 311-324, doi:10.1016/j.cell.2013.06.026 (2013).

34 Jiang, D. *et al.* Alteration in 5-hydroxymethylcytosine-mediated epigenetic regulation leads to Purkinje cell vulnerability in ATM deficiency. *Brain* **138**, 3520-3536, doi:10.1093/brain/awv284 (2015).

35 Wang, Y. & Zhang, Y. Regulation of TET protein stability by calpains. *Cell Rep* **6**, 278-284, doi:10.1016/j.celrep.2013.12.031S2211-1247(13)00789-4[pii] (2014).

36 Moran-Crusio, K. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer cell* **20**, 11-24, doi:10.1016/j.ccr.2011.06.001 (2011).

37 Montalbán-Loro, R. *et al.* TET3 prevents terminal differentiation of adult NSCs by a non-catalytic action at Snrpn. *Nat Commun* **10**, 1726, doi:10.1038/s41467-019-09665-1 (2019).

38 Caldwell, B. A. *et al.* Functionally distinct roles for TET-oxidized 5-methylcytosine bases in somatic reprogramming to pluripotency. *Mol Cell* **81**, 859-869.e858, doi:10.1016/j.molcel.2020.11.045 (2021).

39 Fang, S. *et al.* Tet inactivation disrupts YY1 binding and long-range chromatin interactions during embryonic heart development. *Nat Commun* **10**, 4297, doi:10.1038/s41467-019-12325-z (2019).

40 Lio, C. W. *et al.* Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife* **5**, doi:10.7554/eLife.18290 (2016).

41 Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417-1430, doi:10.1016/j.cell.2012.11.022S0092-8674(12)01407-9 [pii] (2012).

42 Gu, T. *et al.* DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. *Genome Biol* **19**, 88, doi:10.1186/s13059-018-1464-7 (2018).

43 Verma, N. *et al.* TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet* **50**, 83-95, doi:10.1038/s41588-017-0002-y (2018).

44 Greco, C. M. *et al.* DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nat Commun* **7**, 12418, doi:10.1038/ncomms12418 (2016).

45 Godini, R., Lafta, H. Y. & Fallahi, H. Epigenetic modifications in the embryonic and induced pluripotent stem cells. *Gene Expr Patterns* **29**, 1-9, doi:10.1016/j.gep.2018.04.001 (2018).

46 Dawlaty, M. M. *et al.* Loss of tet enzymes compromises proper differentiation of embryonic stem cells. *Dev Cell* **29**, 102-111, doi:10.1016/j.devcel.2014.03.003S1534-5807(14)00155-5 [pii] (2014).

47 Koh, K. P. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200-213, doi:10.1016/j.stem.2011.01.008S1934-5909(11)00009-9 [pii] (2011).

48 Wu, H. & Zhang, Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* **156**, 45-68, doi:10.1016/j.cell.2013.12.019S0092-8674(13)01588-2 [pii] (2014).

49 Langlois, T. *et al.* TET2 deficiency inhibits mesoderm and hematopoietic differentiation in human embryonic stem cells. *Stem Cells* **32**, 2084-2097, doi:10.1002/stem.1718 (2014).

50 Lu, F., Liu, Y., Jiang, L., Yamaguchi, S. & Zhang, Y. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev* **28**, 2103-2119, doi:10.1101/gad.248005.114 (2014).

51 Yang, J. *et al.* Tet Enzymes Regulate Telomere Maintenance and Chromosomal Stability of Mouse ESCs. *Cell Rep* **15**, 1809-1821, doi:10.1016/j.celrep.2016.04.058 (2016).

52 Ficz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398-402, doi:nature10008 [pii]10.1038/nature10008 (2011).

53 Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-538, doi:10.1038/nature04915 (2006).

54 Ma, Z., Swigut, T., Valouev, A., Rada-Iglesias, A. & Wysocka, J. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. *Nat Struct Mol Biol* **18**, 120-127, doi:10.1038/nsmb.2000 (2011).

55 Li, X. *et al.* Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *Proc Natl Acad Sci U S A* **113**, E8267-e8276, doi:10.1073/pnas.1617802113 (2016).

56 Zhu, M. & Zernicka-Goetz, M. Principles of Self-Organization of the Mammalian Embryo. *Cell* **183**, 1467-1478, doi:10.1016/j.cell.2020.11.003 (2020).

57 Zhang, H. T. & Hiiragi, T. Symmetry Breaking in the Mammalian Embryo. *Annu Rev Cell Dev Biol* **34**, 405-426, doi:10.1146/annurev-cellbio-100617-062616 (2018).

58 Muhr, J. & Ackerman, K. M. in *StatPearls* (StatPearls Publishing Copyright © 2021, StatPearls Publishing LLC., 2021).

59 Bardot, E. S. & Hadjantonakis, A. K. Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. *Mech Dev* **163**, 103617, doi:10.1016/j.mod.2020.103617 (2020).

60 Ferretti, E. & Hadjantonakis, A. K. Mesoderm specification and diversification: from single cells to emergent tissues. *Curr Opin Cell Biol* **61**, 110-116, doi:10.1016/j.ceb.2019.07.012 (2019).

61 Pla, P. & Monsoro-Burq, A. H. The neural border: Induction, specification and maturation of the territory that generates neural crest cells. *Dev Biol* **444 Suppl 1**, S36-s46, doi:10.1016/j.ydbio.2018.05.018 (2018).

62 Serrano Nájera, G. & Weijer, C. J. Cellular processes driving gastrulation in the avian embryo. *Mech Dev* **163**, 103624, doi:10.1016/j.mod.2020.103624 (2020).

63 Nowotschin, S. & Hadjantonakis, A. K. Guts and gastrulation: Emergence and convergence of endoderm in the mouse embryo. *Curr Top Dev Biol* **136**, 429-454, doi:10.1016/bs.ctdb.2019.11.012 (2020).

64 Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**, 15-31, doi:10.1242/dev.050849 (2012).

65 Yamaguchi, S., Shen, L., Liu, Y., Sendler, D. & Zhang, Y. Role of Tet1 in erasure of genomic imprinting. *Nature* **504**, 460-464, doi:10.1038/nature12805 (2013).

66 Li, Z. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509-4518, doi:10.1182/blood-2010-12-325241 (2011).

67 Quivoron, C. *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer cell* **20**, 25-38, doi:10.1016/j.ccr.2011.06.003 (2011).

68 Tsukada, Y., Akiyama, T. & Nakayama, K. I. Maternal TET3 is dispensable for embryonic development but is required for neonatal growth. *Sci Rep* **5**, 15876, doi:10.1038/srep15876 (2015).

69 Dawlaty, M. M. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell* **24**, 310-323, doi:10.1016/j.devcel.2012.12.015S1534-5807(12)00589-8 [pii] (2013).

70 Kang, J. *et al.* Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proc Natl Acad Sci U S A* **112**, E4236-4245, doi:10.1073/pnas.1510510112 (2015).

71 Yue, X., Lio, C. J., Samaniego-Castruita, D., Li, X. & Rao, A. Loss of TET2 and TET3 in regulatory T cells unleashes effector function. *Nat Commun* **10**, 2011, doi:10.1038/s41467-019-09541-y (2019).

72 Dai, H. Q. *et al.* TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. *Nature* **538**, 528-532, doi:10.1038/nature20095 (2016).

73 Buijtendijk, M. F. J., Barnett, P. & van den Hoff, M. J. B. Development of the human heart. *Am J Med Genet C Semin Med Genet* **184**, 7-22, doi:10.1002/ajmg.c.31778 (2020).

74 Katano, W., Moriyama, Y., Takeuchi, J. K. & Koshiba-Takeuchi, K. Cardiac septation in heart development and evolution. *Dev Growth Differ* **61**, 114-123, doi:10.1111/dgd.12580 (2019).

75 Quijada, P., Trembley, M. A. & Small, E. M. The Role of the Epicardium During Heart Development and Repair. *Circ Res* **126**, 377-394, doi:10.1161/circresaha.119.315857 (2020).

76 Gilsbach, R. *et al.* Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. *Nat Commun* **9**, 391, doi:10.1038/s41467-017-02762-z (2018).

77 Xu, Y. *et al.* DNA methylation regulates mouse cardiac myofibril gene expression during heart development. *J Biomed Sci* **22**, 88, doi:10.1186/s12929-015-0203-6 (2015).

78 Grunert, M. *et al.* Comparative DNA methylation and gene expression analysis identifies novel genes for structural congenital heart diseases. *Cardiovasc Res* **112**, 464-477, doi:10.1093/cvr/cvw195 (2016).

79 Sheng, W. *et al.* DNA methylation status of NKX2-5, GATA4 and HAND1 in patients with tetralogy of fallot. *BMC Med Genomics* **6**, 46, doi:10.1186/1755-8794-6-46 (2013).

80 Sheng, W. *et al.* Association of promoter methylation statuses of congenital heart defect candidate genes with Tetralogy of Fallot. *J Transl Med* **12**, 31, doi:10.1186/1479-5876-12-31 (2014).

81 Lan, Y. *et al.* TETs Regulate Proepicardial Cell Migration through Extracellular Matrix Organization during Zebrafish Cardiogenesis. *Cell Rep* **26**, 720-732.e724, doi:10.1016/j.celrep.2018.12.076 (2019).

82 Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev* **19**, 1129-1155, doi:10.1101/gad.1303605 (2005).

83 Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370-374, doi:10.1038/nature11925 [pii] (2013).

84 Xu, Y. *et al.* Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell* **151**, 1200-1213, doi:10.1016/j.cell.2012.11.014 (2012).

85 Cong, L. & Zhang, F. Genome engineering using CRISPR-Cas9 system. *Methods Mol Biol* **1239**, 197-217, doi:10.1007/978-1-4939-1862-1_10 (2015).

86 Sternberg, S. H. & Doudna, J. A. Expanding the Biologist's Toolkit with CRISPR-Cas9. *Mol Cell* **58**, 568-574, doi:10.1016/j.molcel.2015.02.032 (2015).

87 Behringer, R., Gertsenstein, M., Nagy, K. V. & Nagy, A. Differentiating Mouse Embryonic Stem Cells into Embryoid Bodies by Hanging-Drop Cultures. *Cold Spring Harb Protoc* **2016**, doi:10.1101/pdb.prot092429 (2016).

88 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).

89 La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494-498, doi:10.1038/s41586-018-0414-6 (2018).

90 Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386, doi:10.1038/nbt.2859 (2014).

91 Firulli, A. B., McFadden, D. G., Lin, Q., Srivastava, D. & Olson, E. N. Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat Genet* **18**, 266-270, doi:10.1038/ng0398-266 (1998).

92 Liu, W. *et al.* Bmp4 signaling is required for outflow-tract septation and branchial-arch artery remodeling. *Proc Natl Acad Sci U S A* **101**, 4489-4494, doi:10.1073/pnas.0308466101 (2004).

93 Niessen, K. *et al.* Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J Cell Biol* **182**, 315-325, doi:10.1083/jcb.200710067 (2008).

94 Huang, Y., Pastor, W. A., Zepeda-Martinez, J. A. & Rao, A. The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. *Nat Protoc* **7**, 1897-1908, doi:10.1038/nprot.2012.103 [pii] (2012).

95 McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501, doi:nbt.1630 [pii]10.1038/nbt.1630 (2010).

96 He, A. *et al.* Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. *Nat Commun* **5**, 4907, doi:10.1038/ncomms5907 (2014).

97 Steimle, J. D. & Moskowitz, I. P. TBX5: A Key Regulator of Heart Development. *Curr Top Dev Biol* **122**, 195-221, doi:10.1016/bs.ctdb.2016.08.008 (2017).

98 Breunig, J. J. *et al.* Ets Factors Regulate Neural Stem Cell Depletion and Gliogenesis in Ras Pathway Glioma. *Cell Rep* **12**, 258-271, doi:10.1016/j.celrep.2015.06.012 (2015).

99 Ciau-Uitz, A., Wang, L., Patient, R. & Liu, F. ETS transcription factors in hematopoietic stem cell development. *Blood Cells Mol Dis* **51**, 248-255, doi:10.1016/j.bcmd.2013.07.010 (2013).

100 Liu, X. S. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233-247.e217, doi:10.1016/j.cell.2016.08.056 (2016).

101 Huang, X. & Wu, S. M. Isolation and functional characterization of pluripotent stem cell-derived cardiac progenitor cells. *Curr Protoc Stem Cell Biol* **Chapter 1**, Unit 1F.10, doi:10.1002/9780470151808.sc01f10s14 (2010).

102 Zhang, Q. *et al.* Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature* **525**, 389-393, doi:10.1038/nature15252 (2015).

103 Lee, M. *et al.* Engineered Split-TET2 Enzyme for Inducible Epigenetic Remodeling. *J Am Chem Soc* **139**, 4659-4662, doi:10.1021/jacs.7b01459 (2017).

104 Li, L. C. & Dahiya, R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**, 1427-1431 (2002).

105 Sun, D. *et al.* MOABS: model based analysis of bisulfite sequencing data. *Genome Biol* **15**, R38, doi:10.1186/gb-2014-15-2-r38 (2014).

106 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

107 Gilsbach, R. *et al.* Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. *Nat Commun* **5**, 5288, doi:10.1038/ncomms6288 (2014).

108 Li, D. & Rozen, R. Maternal folate deficiency affects proliferation, but not apoptosis, in embryonic mouse heart. *J Nutr* **136**, 1774-1778, doi:10.1093/jn/136.7.1774 (2006).

109 Serra-Juhe, C. *et al.* DNA methylation abnormalities in congenital heart disease. *Epigenetics* **10**, 167-177, doi:10.1080/15592294.2014.998536 (2015).

110 Watkins, M. L. & Botto, L. D. Maternal prepregnancy weight and congenital heart defects in offspring. *Epidemiology* **12**, 439-446 (2001).

111 Bogdanovic, O. *et al.* Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat Genet* **48**, 417-426, doi:10.1038/ng.3522 (2016).

112 Krishnan, A. *et al.* A detailed comparison of mouse and human cardiac development. *Pediatr Res* **76**, 500-507, doi:10.1038/pr.2014.128 (2014).

113 Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355-364, doi:10.1038/nature13992 (2014).

114 Dawlaty, M. M. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166-175, doi:S1934-5909(11)00340-7 [pii]10.1016/j.stem.2011.07.010 (2011).

115 Moses, K. A., DeMayo, F., Braun, R. M., Reecy, J. L. & Schwartz, R. J. Embryonic expression of an Nkx2-5/Cre gene using ROSA26 reporter mice. *Genesis* **31**, 176-180 (2001).

116 D'Amato, G. *et al.* Sequential Notch activation regulates ventricular chamber development. *Nat Cell Biol* **18**, 7-20, doi:10.1038/ncb3280 (2016).

117 Luxan, G. *et al.* Mutations in the NOTCH pathway regulator MIB1 cause left ventricular noncompaction cardiomyopathy. *Nat Med* **19**, 193-201, doi:10.1038/nm.3046 (2013).

118 Gregoire, S. *et al.* Essential and unexpected role of Yin Yang 1 to promote mesodermal cardiac differentiation. *Circ Res* **112**, 900-910, doi:10.1161/CIRCRESAHA.113.259259 (2013).

119 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).

120 Weintraub, A. S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573-1588.e1528, doi:10.1016/j.cell.2017.11.008 (2017).

121 Eagen, K. P. Principles of Chromosome Architecture Revealed by Hi-C. *Trends Biochem Sci* **43**, 469-478, doi:10.1016/j.tibs.2018.03.006 (2018).

122 Mumbach, M. R. *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* **13**, 919-922, doi:10.1038/nmeth.3999 (2016).

123 Ay, F., Bailey, T. L. & Noble, W. S. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res* **24**, 999-1011, doi:10.1101/gr.160374.113 (2014).

124 Morikawa, Y., Leach, J. & Martin, J. F. Yin-Yang 1, a new player in early heart development. *Circ Res* **112**, 876-877, doi:10.1161/CIRCRESAHA.113.300947 (2013).

125 Beketaev, I. *et al.* Critical role of YY1 in cardiac morphogenesis. *Dev Dyn* **244**, 669-680, doi:10.1002/dvdy.24263 (2015).

126 Haldipur, P., Dang, D. & Millen, K. J. Embryology. *Handb Clin Neurol* **154**, 29-44, doi:10.1016/b978-0-444-63956-1.00002-3 (2018).

127 Ko, M. *et al.* Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A* **108**, 14566-14571, doi:10.1073/pnas.1112317108 1081112317108 [pii] (2011).

128 Stanley, E. G. *et al.* Efficient Cre-mediated deletion in cardiac progenitor cells conferred by a 3'UTR-ires-Cre allele of the homeobox gene Nkx2-5. *Int J Dev Biol* **46**, 431-439 (2002).

129 Bauer, C. *et al.* Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J Biol Chem* **290**, 4801-4812, doi:10.1074/jbc.M114.605881 (2015).

130 Lu, M. F., Pressman, C., Dyer, R., Johnson, R. L. & Martin, J. F. Function of Rieger syndrome gene in left-right asymmetry and craniofacial development. *Nature* **401**, 276-278, doi:10.1038/45797 (1999).

131 Xiao, Y. *et al.* Hippo Signaling Plays an Essential Role in Cell State Transitions during Cardiac Fibroblast Development. *Dev Cell* **45**, 153-169.e156, doi:10.1016/j.devcel.2018.03.019 (2018).

132 Veland, N. *et al.* The Arginine Methyltransferase PRMT6 Regulates DNA Methylation and Contributes to Global DNA Hypomethylation in Cancer. *Cell Rep* **21**, 3390-3397, doi:10.1016/j.celrep.2017.11.082 (2017).

133 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218, doi:10.1038/nmeth.2688 (2013).

134 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293, doi:10.1126/science.1181369 (2009).

135 Laugwitz, K. L. *et al.* Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. *Nature* **433**, 647-653, doi:10.1038/nature03215 (2005).

136 Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **1**, 4, doi:10.1186/1471-213x-1-4 (2001).

137 Harikrishnan, K. *et al.* Fibulin-1 suppresses endothelial to mesenchymal transition in the proximal outflow tract. *Mech Dev* **136**, 123-132, doi:10.1016/j.mod.2014.12.005 (2015).

138 Xia, M., Luo, W., Jin, H. & Yang, Z. HAND2-mediated epithelial maintenance and integrity in cardiac outflow tract morphogenesis. *Development* **146**, doi:10.1242/dev.177477 (2019).

139 Xiong, H. *et al.* Single-Cell Transcriptomics Reveals Chemotaxis-Mediated Intraorgan Crosstalk During Cardiogenesis. *Circ Res* **125**, 398-410, doi:10.1161/circresaha.119.315243 (2019).

140 Moretti, A. *et al.* Multipotent embryonic isl1+ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell* **127**, 1151-1165, doi:10.1016/j.cell.2006.10.029 (2006).

141 Grison, M., Merkel, U., Kostan, J., Djinović-Carugo, K. & Rief, M. α-Actinin/titin interaction: A dynamic and mechanically stable cluster of bonds in the muscle Z-disk. *Proc Natl Acad Sci U S A* **114**, 1015-1020, doi:10.1073/pnas.1612681114 (2017).

142 Chen, G., Ning, B. & Shi, T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front Genet* **10**, 317, doi:10.3389/fgene.2019.00317 (2019).

143 DeLaughter, D. M. *et al.* Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Dev Cell* **39**, 480-490, doi:10.1016/j.devcel.2016.10.001 (2016).

144 Morgenthau, A. & Frishman, W. H. Genetic Origins of Tetralogy of Fallot. *Cardiol Rev* **26**, 86-92, doi:10.1097/crd.0000000000000170 (2018).

145 Nie, J. *et al.* Post-transcriptional Regulation of Nkx2-5 by RHAU in Heart Development. *Cell Rep* **13**, 723-732, doi:10.1016/j.celrep.2015.09.043 (2015).

146 Gong, H., Lyu, X., Wang, Q., Hu, M. & Zhang, X. Endothelial to mesenchymal transition in the cardiovascular system. *Life Sci* **184**, 95-102, doi:10.1016/j.lfs.2017.07.014 (2017).

147 Jackson, A. O., Zhang, J., Jiang, Z. & Yin, K. Endothelial-to-mesenchymal transition: A novel therapeutic target for cardiovascular diseases. *Trends Cardiovasc Med* **27**, 383-393, doi:10.1016/j.tcm.2017.03.003 (2017).

148 Kovacic, J. C., Mercader, N., Torres, M., Boehm, M. & Fuster, V. Epithelial-to-mesenchymal and endothelial-to-mesenchymal transition: from cardiovascular development to disease. *Circulation* **125**, 1795-1808, doi:10.1161/circulationaha.111.040352 (2012).

149 Izaguirre, M. F. *et al.* Role of E-cadherin in epithelial architecture maintenance. *Cell Commun Adhes* **17**, 1-12, doi:10.3109/15419061003686938 (2010).

150 Loh, C. Y. *et al.* The E-Cadherin and N-Cadherin Switch in Epithelial-to-Mesenchymal Transition: Signaling, Therapeutic Implications, and Challenges. *Cells* **8**, doi:10.3390/cells8101118 (2019).

151 Gonzalez, D. M. & Medici, D. Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal* **7**, re8, doi:10.1126/scisignal.2005189 (2014).

152 Neeb, Z., Lajiness, J. D., Bolanis, E. & Conway, S. J. Cardiac outflow tract anomalies. *Wiley Interdiscip Rev Dev Biol* **2**, 499-530, doi:10.1002/wdev.98 (2013).

153 Saenz, R. B., Beebe, D. K. & Triplett, L. C. Caring for infants with congenital heart disease and their families. *Am Fam Physician* **59**, 1857-1868 (1999).

154 Yuan, X. *et al.* Disruption of spatiotemporal hypoxic signaling causes congenital heart disease in mice. *J Clin Invest* **127**, 2235-2248, doi:10.1172/jci88725 (2017).

155 Tsagaratou, A. *et al.* TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. *Nat Immunol* **18**, 45-53, doi:10.1038/ni.3630 (2017).